Modified Numerals as Post-suppositions

Adrian Brasoveanu, UC Santa Cruz, abrsvn@gmail.com

February 1, 2012

Abstract

The paper provides a compositional account of cumulative readings with non-increasing modified numerals (a.k.a. van Benthem's puzzle), e.g., *Exactly 3 boys saw exactly 5 movies*. The main proposal is that modified numerals make two kinds of semantic contributions. Their asserted / at-issue contribution is a maximization operator that introduces the maximal set of entities that satisfies their restrictor and nuclear scope. The second contribution is a post-supposition, i.e., a cardinality constraint that needs to be satisfied relative to the context that results *after* the at-issue meaning is evaluated. Thus, the interpretation process ends up giving a kind of *pseudo wide scope* to the cardinality post-suppositions contributed by modified numerals.

1 Cumulativity and Modified Numerals

The main goal of the present paper is to provide a compositional account of cumulative readings with non-increasing modified numerals (a.k.a. van Benthem's puzzle, van Benthem 1986), exemplified in (1) below. We focus on *exactly n* modified numerals, but the same problem arises with other non-increasing numerals, e.g., *at most n*, *up to n*, *as many as n*, *maximally n* etc. We restrict our discussion to class B modified numerals, to use the terminology in Nouwen (2010).

(1) Exactly three^x boys saw exactly five^y movies.

(2) **Figure 1**

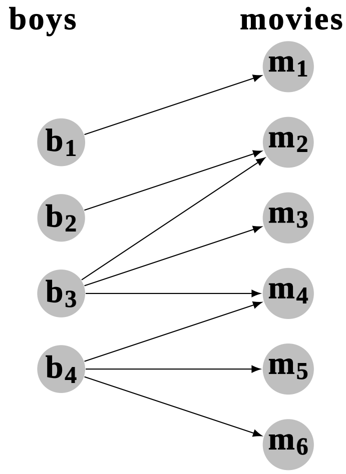
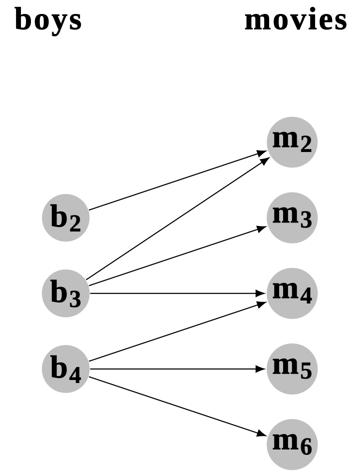


Figure 2



The most familiar reading of sentence (1) (given the common assumption that modified numerals are generalized quantifiers over atomic individuals) is the surface-scope distributive one: there are exactly 3 boys such that each of them saw exactly 5 movies, possibly different from boy to boy.

We are not interested in this reading, although we discuss it briefly later on. Instead, the reading of sentence (1) that we want to capture is the cumulative reading, namely: consider the maximal number of boys that saw a movie and the maximal number of movies seen by a boy; there are 3 such boys and 5 such movies.

Sentence (1) on its cumulative reading could be an exhaustive answer to the question in (3) below, in a situation like the one depicted in Figure 2 above.

(3) How many boys saw how many movies, exactly?

As Krifka (1999), Landman (2000) and Ferreira (2007) observe, the cumulative reading is different from: the maximal number of boys that (between them) saw exactly 5 movies is 3.[1] It is not clear that this is even a possible reading for sentence (1), although it bears some resemblance to its distributive reading.

The situations depicted in Figures 1 and 2 above distinguish between these two readings. Figure 1 is exactly like Figure 2, except for the addition of boy $\mathbf{b_1}$, movie $\mathbf{m_1}$ and the arrow between them symbolizing the seeing relation. The cumulative reading is intuitively false in Figure 1 (4 boys and 6 movies) and true in Figure 2. In contrast, the second ‘reading’ is true in both cases.

The distinction between the cumulative reading and this other ‘reading’ is important for theoretical reasons. Formal systems based on fairly

[1] But see Robaldo (2009) for a different take on the data.

uncontroversial (neo)Montagovian assumptions about semantics and the syntax-semantics interface derive something like it when they attempt to capture the cumulative reading. This is because when we interpret sentence (1) compositionally, the maximality condition contributed by the subject *exactly three boys* takes scope over the maximality condition contributed by the direct object *exactly five movies*. What we want instead is simultaneous global maximization over both subject and direct object, plus interpreting the cardinality requirements (exactly 3 and exactly 5) outside this maximization. But it is not obvious how to get this compositionally.

Given that the cumulative reading and the other 'reading' are not mutually independent, i.e., the non-cumulative reading is weaker than the cumulative reading, we could argue that sentence (1) has only one reading here, namely the weaker, non-cumulative one. This would be parallel to the classical argument that the sentence *Every man loves a woman* does not provide a strong enough argument for quantifier-scope ambiguities, since the wide-scope indefinite 'reading' is simply a strengthening of the weaker, surface-scope one in which the indefinite has narrow scope.

Here are three tests that help us decide whether the non-cumulative 'reading' is really a reading of sentence (1).

First, if the second 'reading' that is not globally maximal is the only one available, we would predict that mentioning the remaining boys and movies is felicitous – at least in situations like the one in Figure 2.

(4) Exactly three^x boys saw exactly five^y movies. #Perhaps there was another boy that saw a different movie, but I didn't notice him.

Second, we can embed sentence (1) under negation and check whether the resulting statement is intuitively true in Figure 1.[2]

(5) It is not true that exactly three^x boys saw exactly five^y movies.[3]
[true in Figure 1, false in Figure 2]

Finally, we can check if sentence (1) could be an exhaustive answer to the question below in both Figure 1 and Figure 2.

(6) How many boys saw how many movies (exactly)?

[2]As an anonymous *Journal of Semantics* reviewer points out, this is really a test for the independence of the cumulative reading and the second 'reading' mentioned above, not a diagnostic that the second 'reading' is not available.

[3]As Philippe Schlenker observes (p.c.), other downward entailing environments could provide more natural settings to test this, e.g., conditionals *If exactly three boys see exactly five movies,...* or the restrictor of universals *In every class in which exactly three boys saw exactly five movies ....*

(7) a. [Figure 2:] Exactly three^x boys saw exactly five^y movies (and no other boy saw no other movie).

b. [Figure 1:] #Exactly three^x boys saw exactly five^y movies and, in addition, Joey saw *Star Wars III*. So, all in all, four boys saw six movies.

The above three tests seem to support the observation that the globally-maximal, cumulative reading is indeed a distinct reading for sentence (1) – and cast doubt on the availability of the second, non-globally maximal 'reading'.

The naturally-occurring examples below from the Corpus of Contemporary American English (COCA, `www.americancorpus.org`) provide additional evidence that globally-maximal cumulative readings with non-increasing quantifiers exist.[4]

(8) [CSX and Norfolk Southern, which haul coal from eastern mines, are making lesser but still sizable investments to maintain or upgrade lines that take a beating from coal trains.]
Due to their size – **up to four locomotives pulling as many as 125 cars** – coal trains wear down tracks more quickly than other cargo haulers.

(9) That's one of the reasons the Alameda County Board of Supervisors approved a proposed upgrade last December by Green Ridge Power LLC, Altamont Power LLC, Sea West Windfarms Inc. and Ventura Pacific Inc. to **replace as many as 1,270 old windmills with up to 187 new ones**.

(10) All measurements made with the unit may be stored in the on-board memory that **will hold as many as 3,000 readings from up to 100 individual probes**.

Further evidence is provided by the example below (also from COCA), where the distributive reading is forced by the addition of the explicit distributor *apiece*. The addition of this explicit distributor points to the fact that the sentence without it can be interpreted cumulatively – and *is* probably interpreted cumulatively by default.

(11) This new organization looks to assist as many as 45 artists per year with loans of up to $3,000 apiece.

Our main proposal is that modified numerals make two kinds of contributions to the meaning of sentences like (1): (*i*) their asserted / at-issue

[4]For more discussion of the less studied construction *as many as n*, see Rett (2010); see Nouwen (2010) and references therein for a detailed discussion of other non-increasing numerals, including *up to*.

contribution is a maximization operator (closely related to the σ operator in Link 1983) that introduces the maximal set of entities that satisfies their restrictor and nuclear scope; (*ii*) the second contribution is a post-supposition, i.e., a cardinality constraint (e.g., exactly three) that needs to be satisfied relative to the *context that results after* the at-issue meaning is evaluated.

Using post-suppositions enables us to interpret the cardinality requirements *globally*, giving them a form of pseudo wide scope, and this is enough: we do not need a global maximality operator. The system is compositional and the maximality conditions are nested, but they are equivalent to a global maximality condition.

The basic insight behind the use of post-suppositions to derive cumulative readings is their pseudo-scopal behavior: interpreting cardinality requirements only *after* the regular at-issue content is interpreted gives them 'wide scope' relative to the maximization operators that are part of that at-issue content. Thus, the syntactic structure of sentence (1) is the normally assumed one and the compositional interpretation procedure has the usual (neo)Montagovian form – but we will effectively interpret (1) as shown in (12) below.

(12) $\sigma xy(\textsc{boy}(x) \wedge \textsc{movie}(y) \wedge \textsc{see}(x, y)) \wedge$
$|y| = 5 \wedge |x| = 3$

This formula has three conjuncts. The first one contains a maximization operator σ that extracts the maximal sum individuals x and y satisfying the (cumulatively-closed) lexical relations in its scope. This conjunct is basically interpreted as: store the maximal number of boys that saw a movie in x and the maximal number of movies seen by a boy in y.

The second and third conjuncts are the post-suppositional cardinality requirements contributed by the two modified numerals. They are post-suppositional in the sense that they are interpreted relative to the context that results after the maximization conjunct is interpreted. That is, they specify the cardinality of the sets of boys and movies that have been previously assigned as values to the variables x and y, respectively (since (12) will be interpreted in a dynamic system, the x and y variables are able to retrieve the values introduced in the first conjunct).

While similar to quantificational items actually taking (syntactic) wide scope, post-suppositional pseudo wide scope has slightly different semantic properties. The differences between class A and class B modified numerals introduced in Nouwen (2010) are due precisely to the fact that, for class A modifiers (which include comparative quantifiers like *fewer than ten guests*), degree quantifiers take actual wide scope along the lines of the analysis in Ferreira (2007),[5] while class B modifiers (which include the modified numerals we are concerned with here) contribute post-suppositional cardinality requirements.

We will work with a modified version of Dynamic Predicate Logic (DPL, Groenendijk & Stokhof 1991) – we slightly change the ontology to add non-atomic individuals. Post-suppositions will be interpreted as cardinality requirements on the maximal non-atomic individuals contributed by modified numerals.

One of the important differences between the present account and the one in Krifka (1999) is conceptual: we take modified numerals to constrain *quantificational* – and not focus – alternatives, where a quantificational alternative is one of the contexts / variable assignments that result after a quantificational expression is interpreted. We therefore predict particular patterns of interaction between the post-suppositions contributed by modified numerals and other quantificational expressions, e.g., modals.

Broader issues related to the place of post-suppositions within the landscape of natural language quantification will not be addressed, except for tentatively suggesting that their pseudo wide scope behavior could be related to the fact that they provide 'wide scope', global level information about the cardinality of quantificational domains. We would therefore expect other items that contribute similarly global information to exhibit a similar scopal behavior.

The paper focuses instead on the inner workings of the proposed post-suppositional mechanism and argues that it makes the right predictions with respect to the interpretation of modified numerals and their interaction with other quantificational expressions.

2 Modified Numerals as Post-suppositions

We work with models that have the same structure as the ones for classical first-order logic (FOL): $\mathfrak{M} = \langle \mathfrak{D}, \mathfrak{I} \rangle$. $\mathfrak{D}$ is the domain of individuals and $\mathfrak{I}$ is the basic interpretation function such that $\mathfrak{I}(R) \subseteq \mathfrak{D}^n$, for any n-ary relation R.

[5]Lucas Champollion (p.c.) brought to my attention the following naturally occurring example of cumulative readings with class A modified numerals:

(1) This book is the product of **more than five hundred hours of interviews with more than two hundred individuals** who participated directly in the events surrounding the financial crisis.
(the first sentence of Andrew Sorkin's "Too Big To Fail")

The difference between our models and the usual FOL ones is that the domain of individuals $\mathfrak{D}$ consists of both atomic individuals and collections / non-atomic individuals. That is, we take $\mathfrak{D}$ to be the power set of a given non-empty set IN of entities: $\mathfrak{D} = \wp^+(\mathsf{IN}) := \wp(\mathsf{IN}) \setminus \{\emptyset\}$.

The sum of two individuals $x \oplus y$ is the union of the sets x and y, e.g., $\{\mathsf{jasper}\} \oplus \{\mathsf{jacob}\} = \{\mathsf{jasper}, \mathsf{jacob}\}$. For a set of atomic / non-atomic individuals X, the sum of the individuals in X (i.e., their union) is $\oplus X$, e.g., $\oplus \{\{\mathsf{jasper}, \mathsf{jacob}\}, \{\mathsf{jacob}\}, \{\mathsf{agatha}\}\} = \{\mathsf{jasper}, \mathsf{jacob}, \mathsf{agatha}\}$.

The part-of relation over individuals $x \leq y$ (x is a part of y) is the partial order induced by inclusion $\subseteq$ over the set $\wp^+(\mathsf{IN})$: $x \leq y := x \subseteq y$. The strict part-of relation is the corresponding order induced by strict inclusion: $x < y := x \subsetneq y$. Atomic individuals are the singleton subsets of IN, identified by means of the predicate $\mathbf{atom}(x)$ defined below.

(13) $\mathbf{atom}(x) := \forall y \leq x(y = x)$

An $\mathfrak{M}$-assignment g is a total function from the set of variables $\mathcal{V}$ to $\mathfrak{D}$. The essence of quantification in FOL is pointwise / variablewise manipulation of variable assignments, abbreviated $h[x]g$. Informally, $h[x]g$ says that assignment h differs from assignment g at most with respect to the value it assigns to the variable x.

(14) $h[x]g :=$ for any variable $v \in \mathcal{V}$, if $v \neq x$, then $h(v) = g(v)$

Note that for any variable $v \in \mathcal{V}$ (x, y, z and so on), the induced binary relation between assignments $h[v]g$ is an equivalence relation: it is reflexive, symmetric and transitive.

The FOL semantic clauses for both universal and existential quantification make use of this kind of assignment manipulation – although the standard formulation of these clauses uses a different kind of notation. Universal and existential FOL quantifiers differ only with respect to *how* the notion of pointwise manipulation of assignments is used: universals look at every assignment h such that $h[x]g$, while existentials only need some assignment h such that $h[x]g$ – where g is the input (contextually-provided) variable assignment relative to which FOL formulas are interpreted.

DPL quantification is the same as FOL quantification. The only difference is that the output assignment h that is the result of the random assignment of value to a variable is preserved and passed on as the input assignment for the next formula. Thus, DPL formulas denote binary relations between input and output contexts.

In particular, atomic formulas for lexical relations are tests, as shown in (15) and (16) below for unary relations (properties) and binary relations, respectively. They require the output context h to be the same as the input context g, i.e., they simply pass on the input context, and they check that h satisfies the lexical relation denoted by P or R. These definitions can be easily generalized to n-ary lexical relations.

(15) $[\![P(x)]\!]^{\langle g,h \rangle} = \mathbb{T}$ iff $g = h$ and $h(x) \in {}^*\mathfrak{I}(P)$

(16) $[\![R(x,y)]\!]^{\langle g,h \rangle} = \mathbb{T}$ iff $g = h$ and $\langle h(x), h(y) \rangle \in {}^{**}\mathfrak{I}(R)$

(17) For any property $\mathbf{P}$ (where $\mathbf{P} = \mathfrak{I}(P)$), the cumulative / sum closure ${}^*\mathbf{P}$ is the smallest set such that $\mathbf{P} \subseteq {}^*\mathbf{P}$ and if $a, a' \in {}^*\mathbf{P}$, then $a \oplus a' \in {}^*\mathbf{P}$.

(18) For any binary relation $\mathbf{R}(= \mathfrak{I}(R))$, the cumulative / sum closure ${}^{**}\mathbf{R}$ is the smallest set such that $\mathbf{R} \subseteq {}^{**}\mathbf{R}$ and if $\langle a, b \rangle, \langle a', b' \rangle \in {}^{**}\mathbf{R}$, then $\langle a \oplus a', b \oplus b' \rangle \in {}^{**}\mathbf{R}$.

The above semantic clauses for atomic formulas are like the corresponding FOL clauses except for: (*i*) the fact that formulas are evaluated relative to a pair of variable assignments (the input assignment g and the output assignment h, which are identical for tests) and (*ii*) the fact that we build cumulativity into the semantic clauses for lexical relations (see Link 1983, Krifka 1986 and Sternefeld 1998 for the original definitions of the operators * and **). We take lexical relations to be cumulatively closed for presentational simplicity: the resulting formulas will be more readable.

Cardinality constraints on the values of variables are also tests, as shown by the semantic clauses below. Cardinality constraints on a variable x simply constrain the number of atoms that are part of the plural individual assigned to x.

(19) $|g(x)|$ is the cardinality of the set of atoms in $g(x)$, i.e.:
$|g(x)| := |\{a \leq g(x) : \mathbf{atom}(a)\}|$

(20) $[\![|x| = n]\!]^{\langle g,h \rangle} = \mathbb{T}$ iff $g = h$ and $|h(x)| = n$

(21) $[\![|x| \leq n]\!]^{\langle g,h \rangle} = \mathbb{T}$ iff $g = h$ and $|h(x)| \leq n$

(22) $[\![|x| \geq n]\!]^{\langle g,h \rangle} = \mathbb{T}$ iff $g = h$ and $|h(x)| \geq n$

Dynamic conjunction and random assignment are defined as usual. In particular, dynamic conjunction is interpreted as relation composition.

(23) $[\![\phi \wedge \psi]\!]^{\langle g,h \rangle} = \mathbb{T}$ iff there is a k such that $[\![\phi]\!]^{\langle g,k \rangle} = \mathbb{T}$ and $[\![\psi]\!]^{\langle k,h \rangle} = \mathbb{T}$

(24) Random assignment:
$[\![[x]]\!]^{\langle g,h \rangle} = \mathbb{T}$ iff $h[x]g$

2.1 Bare Numerals and Singular Indefinites

The format for the translation of singular indefinite articles and bare numerals is provided in (25) below. Intuitively, (25) is true iff n atomic individuals that satisfy ϕ also satisfy ψ. Square brackets [ ] indicate restrictor formulas, round brackets ( ) indicate nuclear scope formulas.

(25) $\exists x[|x| = n \wedge \phi]\ (\psi)$

For singular indefinite articles, n is 1 (which is equivalent to an atomicity requirement). For the bare numeral *two*, n is 2. For the bare numeral *three*, n is 3 etc. Two example translations are provided below.

(26) A^x wolf came in. $\rightsquigarrow \exists x[|x| = 1 \wedge \text{WOLF}(x)]\ (\text{COME-IN}(x))$

(27) Two^x wolves came in. $\rightsquigarrow \exists x[|x| = 2 \wedge \text{WOLF}(x)]\ (\text{COME-IN}(x))$

The compositional translation schema in (25) above is just an abbreviation, as shown in (28) below. We 'decompose' (25) into a flat conjunction of elementary formulas, which are interpreted according to their respective semantic clauses provided above.

(28) $\exists x[|x| = n \wedge \phi]\ (\psi) := [x] \wedge |x| = n \wedge \phi \wedge \psi$

Proper names are interpreted like indefinites, except that their restrictor formula requires the variable to take as its only value the individual that has that name. This is exemplified in (29) below, where JASPER is a non-logical constant denoting the individual Jasper.

(29) $\exists x[x = \text{JASPER}]\ (\phi) := [x] \wedge x = \text{JASPER} \wedge \phi$

Pronouns are indexed with the variable introduced by their antecedent and their translation is that variable itself. We ignore differences between singular and plural pronouns.

For example, the two-sentence discourse in (30) below is compositionally translated as shown in (31a) – and this translation is 'unpacked' in (31b).

(30) A^x wolf came in. It$_x$ bit Jasper^y.

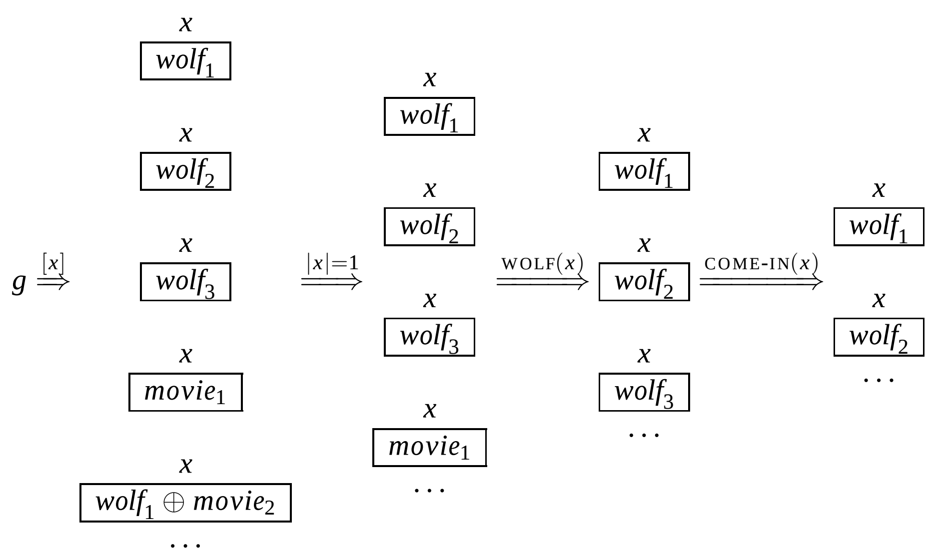
(31) a. $\exists x[|x| = 1 \wedge \text{WOLF}(x)]\ (\text{COME-IN}(x)) \wedge$
$\exists y[y = \text{JASPER}]\ (\text{BITE}(x, y))$

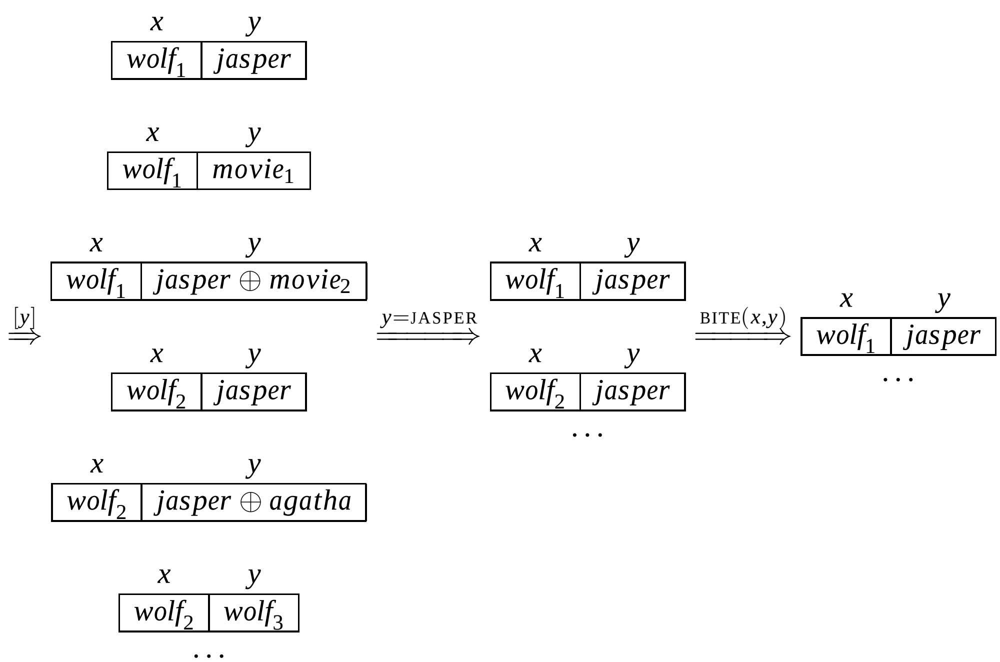
b. $[x] \wedge |x| = 1 \wedge \text{WOLF}(x) \wedge \text{COME-IN}(x) \wedge$
$[y] \wedge y = \text{JASPER} \wedge \text{BITE}(x, y)$

Suppose that our input context is an assignment g that assigns some arbitrary values to all variables. The conjunction of formulas in (31b) above updates this input context as shown in (32) below.

The update in (32) proceeds as follows (recall that the denotations of our formulas are binary *relations*, not functions, between sets of assignments). We first introduce x, i.e., assign it a random value. The result: many contexts / variable assignments that differ from g at most with respect to the value of x and that assign each individual in $\mathfrak{D}$, atomic or not, to x. That is, we now have a graph with many paths. Then, the test $|x| = 1$ eliminates some of the paths in the graph, namely all those paths that end in a context assigning a non-atomic entity to x. The test $\textsc{wolf}(x)$ eliminates further paths in the graph, namely all those that end in a context where x is not assigned a wolf. The test $\textsc{come-in}(x)$ eliminates all the wolves that didn't come in.

We then introduce another variable y that extends the graph in many different ways. The subsequent test $y = \textsc{jasper}$ prunes down the graph by eliminating all contexts that don't assign Jasper to y. Finally, the test $\textsc{bite}(x, y)$ keeps only the output contexts h such that the individual $h(x)$ bit the individual $h(y)$.

(32) 



This interpretation graph is in no way different from the way interpretation proceeds in classical FOL or Discourse Representation Theory (DRT) / File Change Semantics (FCS): such graphs are implicit in their recursive definitions of truth and satisfaction. From now on, we will depict updates by choosing a single, typical path through the graph:

(33) $g \xRightarrow{[x] \wedge |x|=1 \wedge \textsc{wolf}(x) \wedge \textsc{come-in}(x)}$ $\begin{array}{c} x \\ \boxed{wolf_1} \end{array}$

$\xRightarrow{[y] \wedge y=\textsc{jasper} \wedge \textsc{bite}(x,y)}$ $\begin{array}{cc} x & y \\ \boxed{wolf_1} & \boxed{jasper} \end{array}$

The definition of truth below says that a formula is true if there is at least one successful path through the graph / binary relation denoted by ϕ. Again, this is just as in FOL or DRT / FCS.

(34) Truth: a formula ϕ is true relative to an input assignment g iff there is an output assignment h such that $[\![\phi]\!]^{\langle g,h \rangle} = \mathbb{T}$.

Bare numerals are translated and interpreted in a parallel way:

(35) Two^x wolves came in. They$_x$ bit Jasper^y.

(36) a. $\exists x[|x| = 2 \wedge \textsc{wolf}(x)]\ (\textsc{come-in}\,(x)) \wedge$
$\exists y[y = \textsc{jasper}]\ (\textsc{bite}(x, y))$

b. $[x] \wedge |x| = 2 \wedge \textsc{wolf}(x) \wedge \textsc{come-in}\,(x) \wedge$
$[y] \wedge y = \textsc{jasper} \wedge \textsc{bite}(x, y)$

Given that we built cumulativity into the semantic clauses for lexical relations, we get cumulative readings for bare numerals automatically, as shown below. Note that the final output assignment in (39) below encodes a cumulative reading that is true in Figure 1.

(37) Three^x boys saw five^y movies.

(38) a. $\exists x[|x| = 3 \wedge \text{BOY}(x)]\ (\exists y[|y| = 5 \wedge \text{MOVIE}(y)]\ (\text{SEE}(x, y)))$

b. $[x] \wedge |x| = 3 \wedge \text{BOY}(x) \wedge [y] \wedge |y| = 5 \wedge \text{MOVIE}(y) \wedge \text{SEE}(x, y)$

(39) $g \xrightarrow{[x]\wedge|x|=3\wedge\text{BOY}(x)}$

x
$boy_1 \oplus boy_3 \oplus boy_4$

$\xrightarrow{[y]\wedge|y|=5\wedge\text{MOVIE}(y)\wedge\text{SEE}(x,y)}$

x	y
$boy_1 \oplus boy_3 \oplus boy_4$	$movie_1 \oplus movie_2 \oplus movie_3 \oplus movie_4 \oplus movie_5$

2.2 Modified Numerals

We capture the meaning of modified numerals by means of a maximization operator σx closely related to Link's σ operator, which enables us to introduce a maximal plural individual x that satisfies both the restrictor and the nuclear scope formula of a modified numeral.

(40) $[\![\sigma x(\phi)]\!]^{\langle g,h\rangle} = \mathbb{T}$ iff

a. $[\![[x] \wedge \phi]\!]^{\langle g,h\rangle} = \mathbb{T}$

b. there is no h' such that $[\![[x] \wedge \phi]\!]^{\langle g,h'\rangle} = \mathbb{T}$ and $h(x) < h'(x)$

Note that $\sigma x(\phi)$ is a formula and, therefore, denotes a binary relation between variable assignments. It does not denote the maximal plural individual that satisfies the formula ϕ. But after updating an input assignment g with a formula $\sigma x(\phi)$, the output assignment h will indeed assign to x the maximal plural individual satisfying the formula ϕ.[6]

For example, the formula $\sigma x(\text{WOLF}(x))$ introduces the variable x and requires it to store *all* and *only* the atomic individuals satisfying $\text{WOLF}(x)$, i.e., the set of all wolves.

(41) $[\![\sigma x(\text{WOLF}(x))]\!]^{\langle g,h\rangle} = \mathbb{T}$ iff

a. $[\![[x] \wedge \text{WOLF}(x)]\!]^{\langle g,h\rangle} = \mathbb{T}$: we store in x only individuals that satisfy $\text{WOLF}(x)$, i.e., x stores only wolves

b. there is no h' such that $[\![[x] \wedge \text{WOLF}(x)]\!]^{\langle g,h'\rangle} = \mathbb{T}$ and $h(x) < h'(x)$: there is no way to store more atoms in x and still satisfy $\text{WOLF}(x)$, i.e., x stores all the wolves (maximality)

[6]The clause in (40) is actually more general and allows for cases in which there is no single maximal individual, i.e., there is no supremum. In that case, an output assignment h will assign to x one of the maximal plural individuals satisfying the formula ϕ.

We can now provide a preliminary translation for modified numerals (to be modified in important ways):

(42) *exactly n* $\exists|x| = n[\phi]\ (\psi) := \sigma x(\phi \wedge \psi) \wedge |x| = n$

(43) *at most n* $\exists|x| \leq n[\phi]\ (\psi) := \sigma x(\phi \wedge \psi) \wedge |x| \leq n$

(44) *at least n* $\exists|x| \geq n[\phi]\ (\psi) := \sigma x(\phi \wedge \psi) \wedge |x| \geq n$

For example, the sentence in (45) below is compositionally translated as shown in (46a), which is just an abbreviation of the formula in (46b). Intuitively, we store in x all the atomic entities that are wolves and that came in; then, we test that there are 3 such atomic entities.

(45) Exactly three^x wolves came in.

(46) a. $\exists|x| = 3[\text{WOLF}(x)]\ (\text{COME-IN}(x))$
b. $\sigma x(\text{WOLF}(x) \wedge \text{COME-IN}(x)) \wedge |x| = 3$

We obtain similar translations for other modified numerals.

(47) At most three^x wolves came in.
a. $\exists|x| \leq 3[\text{WOLF}(x)]\ (\text{COME-IN}(x))$
b. $\sigma x(\text{WOLF}(x) \wedge \text{COME-IN}(x)) \wedge |x| \leq 3$

(48) At least three^x wolves came in.
a. $\exists|x| \geq 3[\text{WOLF}(x)]\ (\text{COME-IN}(x))$
b. $\sigma x(\text{WOLF}(x) \wedge \text{COME-IN}(x)) \wedge |x| \geq 3$

We can further elaborate on all the above sentences with *They*$_x$ *bit Jasper*^{y} and derive intuitively correct truth conditions for the resulting discourses.

For example, for sentence (48) above, we correctly capture the maximality of cross-sentential anaphora to the modified numeral: at least three wolves came in and all the wolves that came in bit Jasper.

(49) a. $\exists|x| \geq 3[\text{WOLF}(x)]\ (\text{COME-IN}(x)) \wedge$
$\exists y[y = \text{JASPER}]\ (\text{BITE}(x, y))$
b. $\sigma x(\text{WOLF}(x) \wedge \text{COME-IN}(x)) \wedge |x| \geq 3 \wedge$
$[y] \wedge y = \text{JASPER} \wedge \text{BITE}(x, y)$

But we derive incorrect truth conditions for sentence (1). As the formulas in (50) below show, we do not derive the cumulative reading, true only in Figure 2. Instead, we derive the 'reading' true in both Figure 1 and Figure 2: the maximal number of boys that saw exactly 5 movies is 3. This is because the cardinality requirement $|y| = 5$ contributed by the direct object is in the scope of the maximization operator $\sigma x(\dots)$ contributed by the subject.

(50) a. $\exists|x| = 3[\text{BOY}(x)]\ (\exists|y| = 5[\text{MOVIE}(y)]\ (\text{SEE}(x, y)))$

b. $\sigma x(\textsc{boy}(x) \wedge \sigma y(\textsc{movie}(y) \wedge \textsc{see}(x,y)) \wedge |y| = 5) \wedge |x| = 3$

What we want is a translation that places the cardinality requirement $|y| = 5$ contributed by the direct object outside the scope of the maximization operator contributed by the subject:

(51) $\sigma x(\textsc{boy}(x) \wedge \sigma y(\textsc{movie}(y) \wedge \textsc{see}(x,y))) \wedge |y| = 5 \wedge |x| = 3$

The formula in (51) above is equivalent to the one in (52) below. That is, if the cardinality requirement $|y| = 5$ is somehow scoped out, the two nested maximization operators $\sigma x(\ldots \sigma y(\ldots))$ contributed by the subject and the direct object are equivalent to a *global* maximization operator $\sigma xy(\ldots)$, defined in (53) below.

(52) $\sigma xy(\textsc{boy}(x) \wedge \textsc{movie}(y) \wedge \textsc{see}(x,y)) \wedge |y| = 5 \wedge |x| = 3$

(53) $[\![\sigma xy(\phi)]\!]^{\langle g,h \rangle} = \mathbb{T}$ iff

a. $[\![[x] \wedge [y] \wedge \phi]\!]^{\langle g,h \rangle} = \mathbb{T}$

b. there is no h' such that $[\![[x] \wedge [y] \wedge \phi]\!]^{\langle g,h' \rangle} = \mathbb{T}$ and $h(x) < h'(x)$ or $h(y) < h'(y)$

The fact that the two nested maximization operators $\sigma x(\ldots \sigma y(\ldots))$ can be reduced to the global maximization operator $\sigma xy(\ldots)$ in this case follows from the semantics of the σ operators and the fact that all the formulas in their scope, i.e., $\textsc{boy}(x)$, $\textsc{movie}(y)$ and $\textsc{see}(x,y)$, are cumulatively-closed lexical relations.[7]

The formula in (52) closely reflects the intuitive characterization of the cumulative reading of sentence (1): a global maximization operator over both boys and movies is followed by the relevant cardinality requirements. That is, we introduce the maximal set x of boys that saw a movie and the maximal set y of movies seen by a boy and check that there are 5 such movies and 3 such boys.

2.3 Post-suppositions

To be able to compositionally derive such a representation, we will take cardinality requirements to be part of a dimension of meaning separate from the asserted / at-issue meaning – but closely integrated with it. We will take them to be *post-suppositions*, i.e., tests on the output context, as opposed to presuppositions, which are tests on the input context. See Lauer (2009) for another use of the same notion and Farkas (2002) and Constant (2006) for related types of post-assertion constraints on output contexts.

[7] The equivalence holds for a larger class of formulas, not only lexical relations, e.g., the equivalence also holds if there are random assignment formulas $[z]$ in the scope of the σ operators.

Post-suppositions are formulas introduced at certain points in the interpretation that are passed on from local context to local context and that need to be satisfied only globally, relative to the final output context.

Thus, we enrich our notion of evaluation contexts: a context is an assignment g indexed with a set of tests ζ, represented as $g[\zeta]$. That is, output contexts will not store only the values of the previously introduced variables, but also constraints on these values. This is parallel to the way in which presuppositions anaphorically retrieve the values of certain variables stored in the input context and, in addition, check that these values satisfy various constraints.

All the operators above are interpreted in the same way except that, if the input context g is indexed with a set of tests ζ, this set is passed on to the output context h. The interpretation function is not simply $[\![\cdot]\!]^{\langle g,h\rangle}$, but $[\![\cdot]\!]^{\langle g[\zeta],h[\zeta']\rangle}$, where ζ and ζ' are sets of tests and $\zeta \subseteq \zeta'$.

We mark a formula ϕ as a post-supposition by superscripting it, as shown below.

(54) $[\![^{\phi}]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff ϕ is a test, $g = h$ and $\zeta' = \zeta \cup \{\phi\}$.[8]

A post-suppositional formula does not update the input assignment g in any way. It is simply added to the input set of tests ζ.

These tests are post-suppositional in the sense that they are required to be true relative to the final output context – which grants them something very similar to widest scope. This is formalized by means of the definition of truth below.

(55) Truth: a formula ϕ is true relative to an input context $g[\emptyset]$, where $\emptyset$ is the empty set of tests, iff there is an output assignment h and a (possibly empty) set of tests $\{\psi_1, \ldots, \psi_m\}$ such that
 a. $[\![\phi]\!]^{\langle g[\emptyset],h[\{\psi_1,\ldots,\psi_m\}]\rangle} = \mathbb{T}$ and
 b. $[\![\psi_1 \wedge \ldots \wedge \psi_m]\!]^{\langle h[\emptyset],h[\emptyset]\rangle} = \mathbb{T}$.

The definition of truth treats the formulas $\psi_1, \ldots, \psi_m$ as post-suppositions, i.e., as tests performed on the final output context h – as opposed to presuppositions, i.e., tests performed on input contexts.

Recall that dynamic conjunction over tests has the same properties as classical static conjunction – in particular, it is commutative. So, it does not matter in which order we conjoin the tests $\psi_1, \ldots, \psi_m$ when we check that the final output assignment h satisfies them.

The entire recursive definition of truth and satisfaction needs to be reformulated in terms of assignments indexed with sets of tests $g[\zeta]$

[8] ϕ is a test iff for any assignments g and h and any sets of formulas ζ and ζ', if $[\![\phi]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$, then $g = h$ and $\zeta = \zeta'$.

rather than simply assignments g. The relevant definitions are provided below. They are minimally different from the DPL-style ones provided above – we simply index input and output assignments with sets of tests ζ and ζ' and require that $\zeta = \zeta'$ for all non-post-suppositional / at-issue updates.

(56) $[\![P(x)]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and $h(x) \in {}^*\Im(P)$

(57) $[\![R(x,y)]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and $\langle h(x), h(y)\rangle \in {}^{**}\Im(R)$

(58) $[\![|x| = n]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and $|h(x)| = n$

(59) $[\![|x| \leq n]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and $|h(x)| \leq n$

(60) $[\![|x| \geq n]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and $|h(x)| \geq n$

(61) $[\![\phi \wedge \psi]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff there is a k and a ζ'' such that $[\![\phi]\!]^{\langle g[\zeta],k[\zeta'']\rangle} = \mathbb{T}$ and $[\![\psi]\!]^{\langle k[\zeta''],h[\zeta']\rangle} = \mathbb{T}$

(62) $[\![[x]]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $h[x]g$ and $\zeta = \zeta'$

(63) $[\![\sigma x(\phi)]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff

a. $[\![[x] \wedge \phi]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$

b. there is no h' such that $[\![[x] \wedge \phi]\!]^{\langle g[\zeta],h'[\zeta']\rangle} = \mathbb{T}$ and $h(x) < h'(x)$

(64) $[\![\sigma xy(\phi)]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff

a. $[\![[x] \wedge [y] \wedge \phi]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$

b. there is no h' such that $[\![[x] \wedge [y] \wedge \phi]\!]^{\langle g[\zeta],h'[\zeta']\rangle} = \mathbb{T}$ and $h(x) < h'(x)$ or $h(y) < h'(y)$

Modified numerals are interpreted as before, except that the cardinality requirements are now post-suppositional.

(65) *exactly n* $\quad \exists^{|x|=n}[\phi]\ (\psi) := \sigma x(\phi \wedge \psi) \wedge {}^{|x|=n}$

(66) *at most n* $\quad \exists^{|x|\leq n}[\phi]\ (\psi) := \sigma x(\phi \wedge \psi) \wedge {}^{|x|\leq n}$

(67) *at least n* $\quad \exists^{|x|\geq n}[\phi]\ (\psi) := \sigma x(\phi \wedge \psi) \wedge {}^{|x|\geq n}$

That is, numeral modifiers *exactly*, *at most*, *at least* etc. can be thought of as functions that take a bare numeral as their argument and introduce: (*i*) a maximization operator σ that scopes over the random assignment and the restrictor and nuclear scope formulas and (*ii*) a post-supposition that consists of the cardinality requirement (the bare numeral contributes this cardinality requirement to the regular / at-issue content).

Now, the translation of sentence (1) derives the intuitively correct cumulative truth conditions, as shown in (68) below.

(68) a. $\exists^{|x|=3}[\textsc{boy}(x)]\ (\exists^{|y|=5}[\textsc{movie}(y)]\ (\textsc{see}(x,y)))$
b. $\sigma x(\textsc{boy}(x) \wedge \sigma y(\textsc{movie}(y) \wedge \textsc{see}(x,y)) \wedge {}^{|y|=5}) \wedge {}^{|x|=3}$
c. $\sigma x(\textsc{boy}(x) \wedge \sigma y(\textsc{movie}(y) \wedge \textsc{see}(x,y))) \wedge {}^{|y|=5} \wedge {}^{|x|=3}$
d. $\sigma xy(\textsc{boy}(x) \wedge \textsc{movie}(y) \wedge \textsc{see}(x,y)) \wedge {}^{|y|=5} \wedge {}^{|x|=3}$
e. $\sigma xy(\textsc{boy}(x) \wedge \textsc{movie}(y) \wedge \textsc{see}(x,y)) \wedge |y| = 5 \wedge |x| = 3$

All the formulas in (68) are truth-conditionally equivalent (given the definition of truth in (55)). Let us examine them in turn.

The formula in (68a) is what we derive if we follow our compositional translation schemas. The next formula (68b) unpacks (68a) based on the abbreviations defined for each of the translation schemas. These two formulas follow more or less immediately from the basic definitions. The following ones however exploit the more interesting properties of the system we have set up.

Formula (68c) is just like (68b) except that the post-supposition ${}^{|y|=5}$ contributed by the direct object is extracted from the scope of the σx operator contributed by the subject and is placed at the 'top level'. Formulas (68c) and (68b) are equivalent (with respect to both truth conditions and context-change potential) because post-suppositions are simply collected and passed on from local context to local context and are required to be satisfied only relative to the final output context.

Formula (68d) is just like (68c) except that the nested maximization operators have been replaced with a single global operator. As we already observed in the previous subsection, the equivalence of these two formulas (with respect to both truth conditions and context-change potential) follows from the semantics of the σ operators and the fact that they only have cumulatively-closed lexical relations in their scope.

Finally, formula (68e) is just like (68d) except that the update-final post-suppositions ${}^{|y|=5}$ and ${}^{|x|=3}$ are converted into at-issue tests $|y| = 5$ and $|x| = 3$. The truth-conditional equivalence of the two formulas (note that they do not have the same context-change potential) can be derived from the definition of truth as follows.

Given an input context $g[\emptyset]$, assume that the formula (68d) is true relative to $g[\emptyset]$. By the definition of truth in (55) above, (68d) is true iff:

(69) a. there is an output context $h[\{|y| = 5, |x| = 3\}]$ such that $[\![(68d)]\!]^{\langle g[\emptyset], h[\{|y|=5,|x|=3\}]\rangle} = \mathbb{T}$
b. $h[\emptyset]$ satisfies both tests in the set $\{|y| = 5, |x| = 3\}$

Given the semantic clauses for the σ operator, dynamic conjunction and post-suppositions in (64), (61) and (54) above (respectively), we can further simplify (69a) as follows:

(70) $[\![(68d)]\!]^{\langle g[\emptyset],h[\{|y|=5,|x|=3\}]\rangle} = \mathbb{T}$ iff
$[\![\sigma xy(\text{BOY}(x) \wedge \text{MOVIE}(y) \wedge \text{SEE}(x,y))]\!]^{\langle g[\emptyset],h[\emptyset]\rangle} = \mathbb{T}$

From (70), it follows that $h[\emptyset]$ is the output context that obtains after we interpret the maximization update $\sigma xy(\text{BOY}(x) \wedge \text{MOVIE}(y) \wedge \text{SEE}(x,y))$ relative to $g[\emptyset]$. In addition, (69b) states that $h[\emptyset]$ satisfies the tests $|y| = 5$ and $|x| = 3$. But these are exactly the two statements needed to establish that $[\![(68e)]\!]^{\langle g[\emptyset],h[\emptyset]\rangle} = \mathbb{T}$. Hence, we have that (68e) is also true relative to the input context $g[\emptyset]$.

We have just sketched the proof for: if (68d) is true relative to an input context $g[\emptyset]$, then (68e) is also true relative to $g[\emptyset]$. The proof for the other direction is similar and we omit it.

Importantly, the analysis does not predict that sentence (1) is equivalent to the minimally different sentences below, in which the modified numerals are replaced by the corresponding bare numerals (one at a time):

(71) Three^x boys saw exactly five^y movies.

(72) Exactly three^x boys saw five^y movies.

This is because modified numerals differ from bare numerals in two ways. On one hand, the cardinality requirement is a regular at-issue update for bare numerals and a post-supposition for modified numerals. On the other hand, modified numerals contribute a maximization operator, while bare numerals do not. We attributed both differences to the presence of the modifier *exactly* (or *at least*, *at most* etc.).

Let us consider more closely the latter difference between modified and bare numerals. Independent justification for this maximization operator is provided by the contrast between the felicitous bare numeral *two* and the infelicitous modified numeral *at least two* in the discourse below (from Umbach 2006; see also Szabolcsi 1997, de Swart 1999 and Krifka 1999).

(73) a. { Two / #At least two } boys were selling coke.
b. They were wearing black leather jackets.
c. Perhaps there were others also selling coke, but I didn't notice.

The variable introduced by the bare numeral *two* has different sets of two boys as values in different output contexts if more than two boys were selling coke. In the present system, this referential indeterminacy is captured by the fact that the output contexts / variable assignments obtained after the update with a bare numeral might assign different plural individuals to the variable contributed by the bare numeral. For example, if three boys boy_1, boy_2 and boy_3 were selling coke, there are three possible output contexts after the update with the bare numeral, as shown below.

(74) $g \xrightarrow{[x]\wedge|x|=2\wedge\textsc{boy}(x)\wedge\textsc{sell-coke}(x)} \begin{array}{c} x \\ \boxed{boy_1 \oplus boy_2} \\ \\ x \\ \boxed{boy_1 \oplus boy_3} \\ \\ x \\ \boxed{boy_2 \oplus boy_3} \end{array}$

In contrast, the variable introduced by *at least two* has only one possible value: the set of all boys who were selling coke. Determined reference means just this: in any given world, all output contexts obtained after the update with a modified numeral assign the same value to the variable contributed by the modified numeral. In the situation we are considering, this is the set containing the three boys that were selling coke.

(75) $g \xrightarrow{\sigma x(\textsc{boy}(x)\wedge\textsc{sell-coke}(x))\wedge^{|x|\geq 2}} \begin{array}{c} x \\ \boxed{boy_1 \oplus boy_2 \oplus boy_3} \end{array}$

The *perhaps* continuation might not be the best diagnostic, since such continuations are felicitous in many examples from COCA:[9]

(76) Areva, the world's biggest nuclear power plant construction company, announced it would build **at least two, and perhaps six**, EPR nuclear reactors in India.

(77) An American platoon surprised an armed Taliban column on a forested ridgeline at night, and killed **at least 13 insurgents, and perhaps many more**, with rifles, machine guns, Claymore mines, hand grenades and a knife.

(78) **At least one and perhaps two** of the first four Rotarians were Masons.

(79) Up until the last 20 years, vaccines contained **at least 200 and perhaps more than 3,000** antigens.

(80) **At least 4,000 people, and perhaps as many as 6,500**, were killed.

[9] See also the discussion in Geurts & Nouwen (2007) of examples like *John invited at least two boys, namely Pete and Billy*, where it seems that we can refer to witnesses independently of / in addition to the maximal witness set introduced by *at least* (I am indebted to Rick Nouwen, p.c., for discussion of this point). It might be that modified numerals introduce multiple discourse referents, but we will ignore this issue here.

But, at the same time, there are plenty examples in COCA in which plural anaphora to modified numerals is clearly maximal and which do actually provide additional support for the proposal that modified numerals contribute a maximization operator.

(81) *At least*:

a. There were **at least 40 shots**. **They** were single shots but fairly close together.

b. Most Europeans speak **at least two languages** and **they** speak them well by the time they're out of school.

(82) *At most*:

a. Today's main organization, the FARC, had **at most 500 soldiers** – and **they** prowled the most isolated areas of the country.

(83) *Up to*:

a. Program **up to 37 alerts**; **they**'ll reset automatically at midnight.

b. So you can submit **up to three entries**. **They**'ll be judged on meaning, naturalness of syntax, originality and overall elegance.

(84) *As many as*:

a. Out of some 14,000 wildebeests, **as many as 3,000** behaved as permanent residents. **They** could be found in specific areas.

b. There are, depending on how you count them, perhaps **as many as 720 national laboratories**. **They** have collective budgets of more than twenty billion dollars.

It is possible that the contrast between the referential indeterminacy of bare numerals and the referential determinacy of modified numerals (due to maximization) might be the reason behind the fact that modified numerals, unlike bare numerals / indefinites, do not trigger scalar implicatures.[10]

Finally, note that the bipartite structure of the meaning for modified numerals, i.e.,

[10]Under the (unconventional) assumption that scalar implicatures are just a way to pragmatically resolve the referential indeterminacy associated with bare numerals and singular indefinites, the presence of maximization would enable us to explain the fact that modified numerals do not trigger scalar implicatures despite the fact that they can easily be associated with Horn scales – e.g., in upward-entailing contexts, *at least three* is less informative than *at least four*, which is less informative than *at least five* etc. We leave the investigation of this conjecture for a future occasion.

(*i*) first, introduce a plural individual that is maximal relative to the part-of ordering $\leq$

(*ii*) then, specify by means of a post-supposition how this plural individual is related to a specific point on a (measurement) scale associated with the same ordering (a specific cardinality, i.e., a specific number of $\leq$-atoms)

can be generalized to other ontological domains (besides individuals) that come with different kinds of orderings and associated measurement scales (this is related to the suggestion in Nouwen 2010 that class B modifiers are about maxima / minima).[11]

The basic structure of the update should be the same: first, introduce a variable of the appropriate type and assign it a value that is maximal relative to the contextually-provided ordering; then, specify how that value is related to a specific point on the contextually-provided scale associated with the ordering.

The range of objects, orderings and scales that class B modifiers can be used to refer to is fairly broad, as the *at least* examples below from COCA indicate. The idea is that as long as we have variables for objects of arbitrary static types (there are dynamic systems that allow this), we do not need to invoke additional machinery, e.g., focus alternatives, to capture the range of uses of class B modifiers.

(85) Finally, the Americans sent a convoy of soldiers speeding into the valley to support or save their allies or **at least secure** the dead.

(86) The themes and expressions of **at least some** of the sculptures are likely to be influenced by our conversations.

(87) I would argue that it is the condition of the works' ephemerality that enables or **at least allows** for these " truth " tales to be spun.

(88) It is **at least possible** that they – some of them – are serious about finding a peaceful way out.

(89) Their forces in the K-G Pass seemed crippled by the losses, **at least temporarily**.

(90) Shouldn't we **at least occasionally** think about how we want to leave our lives?

(91) These factors **at least partly** explain the amount of criticism directed at Israel.

(92) "**At least he's here**," Jan says. "**At least there's that**."

[11] I am grateful to Lucas Champollion and an anonymous *Journal of Semantics* reviewer for emphasizing this point.

In sum, the proposed analysis of modified numerals makes use of three formal ingredients: (*i*) post-suppositions and their unusual 'scoping' behavior, (*ii*) a dynamic interpretation procedure that preserves the output context so that post-suppositions can be checked against it (in contrast, static semantics 'flushes' the output context once quantifiers are interpreted and resets everything to the input context) and (*iii*) maximization operators that store maximal plural individuals relative to certain variables.

Post-suppositions are the only ingredient requiring independent motivation – the other two are fairly uncontroversial and commonly used in the semantic literature.[12]

The following section (section 3) on the interaction between modified numerals and modals provides independent motivation for post-suppositions. But before we turn to this, we discuss the distributive readings of modified numerals and the scopal interactions between them and distributive quantifiers like *every+NP*.

2.4 Distributive Readings

Up until now, we have shown how cumulative readings for both bare and modified numerals can be compositionally captured. But both kinds of numerals can also have distributive readings. Moreover, such distributive readings are the default ones for universal quantifiers like *every+NP* or *each+NP*.

A simple way to capture distributive readings is to add a distributive operator δx, needed not only for distributively-interpreted bare or modified numerals, but also for the interpretation of universal quantification, as shown in (93) below.

(93) $\forall x[\phi]\ (\psi) := \sigma x(\phi) \wedge \delta x(\psi)$

A universal quantifier introduces the set of all individuals x that satisfy the restrictor ϕ – by means of $\sigma x(\phi)$ – and then checks that *each* of these individuals also satisfies the nuclear scope ψ – by means of $\delta x(\psi)$.

We define the distributivity operator δx in two steps.

Let us first ignore post-suppositions and define the basic notion of distributivity.

[12]This might sound as an overstatement with respect to the second ingredient, but situation-based E- / D-type accounts of donkey anaphora simply encapsulate the dynamics of interpretation in the rules for extending / updating minimal situations. See Brasoveanu (2008) and Dekker (2010) for a more detailed discussion of the common insights behind both dynamic and situation semantics.

(94) $[\![\delta x(\phi)]\!]^{\langle g,h\rangle} = \mathbb{T}$ iff $g = h$ and for all atoms $a \leq g(x)$, if we let g' be such that $g'[x]g$ and $g'(x) = a$, then there is a k such that $[\![\phi]\!]^{\langle g',k\rangle} = \mathbb{T}$

Informally, we distributively interpret a formula ϕ relative to a plural individual $g(x)$ by checking that ϕ is satisfied by each atom a that is part of $g(x)$.

The definition above does this by temporarily reassigning the variable x so that it stores each atom a relative to the variable x, one atom at a time. Relative to this temporary new assignment g', we check that a further update with the formula ϕ is possible, i.e., we check that there is at least one possible output assignment k such that ϕ is true relative to the pair of assignments $\langle g', k\rangle$.

After we loop through all atoms a that are part of $g(x)$ and check that, for each of them, a further update with ϕ is possible, we know that the plural individual $g(x)$ distributively satisfies ϕ, so we simply pass on the original input assignment g. Thus, distributive formulas of the form $\delta x(\phi)$ are tests.

Taking quantificational distributivity to contribute a test is not a necessary feature of all possible dynamic systems. In fact, it is mistaken to do so given that subsequent anaphora to universal quantifiers and indefinites in their scope is possible, as the example of quantificational subordination in (95) below (from Karttunen 1976) shows.

(95) a. Harvey^x courts a^y woman at every^z convention.
b. She$_y$ always$_z$ comes to the banquet$_z$ with him$_x$.
c. The$_y$ girl is usually$_z^{z'}$ also very pretty.

But given our DPL-style system in which output contexts are single assignments, taking distributivity to contribute a test is more or less a forced choice.[13]

We return to the issue of distributivity when we discuss cumulative readings for universal quantifiers in section 4. The important point here is that it is natural to interpret distributive updates as tests, just as it is natural to interpret distributive quantifiers as tests in DRT and FCS.

The second step in our definition of distributivity δx is to bring post-suppositions into the picture.

The addition of post-suppositions makes necessary two modifications to the definition above. The first one is trivial: just as the distributivity operator is a test relative to the input variable assignment g, i.e.,

[13]For various ways of extending DPL to keep track of the quantificational dependencies introduced by (and within the scope of) distributive quantification, see van den Berg (1996), Nouwen (2003), Brasoveanu (2007) and Dekker (2008) among others.

we require that $g = h$, the distributivity operator is also a test relative to the input set of post-suppositions ζ, i.e., we require $\zeta = \zeta'$.

(96) $[\![\delta x(\phi)]\!]^{\langle g[\zeta], h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and for all atoms $a \leq g(x)$, if we let g' be such that $g'[x]g$ and $g'(x) = a$, then there is a k and a (possibly empty) set of tests $\{\psi_1, \ldots, \psi_m\}$ such that

a. $[\![\phi]\!]^{\langle g'[\zeta], k[\zeta \cup \{\psi_1, \ldots, \psi_m\}]\rangle} = \mathbb{T}$ and

b. $[\![\psi_1 \wedge \ldots \wedge \psi_m]\!]^{\langle k[\zeta], k[\zeta]\rangle} = \mathbb{T}$

The second modification is less obvious: we let the distributivity operator δx discharge all the post-suppositions $\{\psi_1, \ldots, \psi_m\}$ contributed by the formula ϕ in its scope, as shown in (96b) above.

The reason for this is that distributivity is externally static, i.e., it is a test relative to the input variable assignment g: whatever new variables / discourse referents are introduced in the scope of the operator δx, they are not accessible outside its scope since the output assignment h is the same as the input assignment g.

Therefore, any post-suppositional cardinality requirements introduced within the scope of the distributivity operator δx have to be discharged / satisfied locally. If they were passed on to the output context h, they would try to access variables whose values are not available in that context because their values were both introduced and 'flushed' within the scope of δx.

In a sense, distributivity operators are to post-suppositions what clausal boundaries are to movement in syntax: they mark locality domains / barriers. Thus, just like presuppositions or scalar implicatures in theories like Chierchia et al. (2009), post-suppositions are not always satisfied globally, but can be satisfied / discharged at intermediate points in the semantic composition, i.e., in more local output contexts.

This is not unexpected. We work with quantificational alternatives, i.e., with contexts / assignments that are the result of interpreting quantificational expressions, and not with focus alternatives, as Krifka (1999). So we expect various *quantificational* operators (universals, modals, attitude verbs, negation etc.) that are interpreted as tests in DRT / FCS / DPL style systems to block the 'projection' of post-suppositions and discharge them locally, in their scope.

The distributive operator δx enables us to capture both the scopal interactions between modified numerals and universal quantifiers and the distributive readings of the modified numerals themselves.

Let us examine the latter first. As we observed at the very beginning of the paper, sentence (1) has a distributive reading, which is probably the most familiar one in the formal semantics literature. To derive this reading, we assume that modified numerals can optionally be distributive, i.e., can optionally have a δ operator over their nuclear scope, as shown in (97) below.

(97) $\exists^{|x|=n}[\phi]\ \delta(\psi) := \sigma x(\phi \wedge \delta x(\psi)) \wedge\ ^{|x|=n}$

The surface-scope distributive reading for sentence (1) is provided in (98) below; the three formulas in (98) are truth-conditionally equivalent. Note that the embedded distributive operator δy contributed by the direct object is semantically vacuous in this case, so it is omitted in (98c) to improve readability. Also, note that the post-supposition $^{|y|=5}$ contributed by the direct object is trapped in the scope of the operator δx contributed by the subject, so it is locally discharged / satisfied. That is, it can be locally replaced with a non-post-suppositional test $|y| = 5$, as shown in (98c).

(98) a. $\exists^{|x|=3}[\text{BOY}(x)]\ \delta(\exists^{|y|=5}[\text{MOVIE}(y)]\ \delta(\text{SEE}(x,y)))$

b. $\sigma x(\text{BOY}(x) \wedge$
$\quad \delta x(\sigma y(\text{MOVIE}(y) \wedge \delta y(\text{SEE}(x,y))) \wedge\ ^{|y|=5})) \wedge$
$^{|x|=3}$

c. $\sigma x(\text{BOY}(x) \wedge$
$\quad \delta x(\sigma y(\text{MOVIE}(y) \wedge \text{SEE}(x,y)) \wedge |y| = 5)) \wedge$
$|x| = 3$

The formula in (98c) is intuitively interpreted as follows. First, introduce the set of all boys x such that, when we take each boy one at a time, there are exactly 5 movies y that each of them saw. That is, store in x all the boys such that each of them saw exactly 5 movies. Then, check that x stores exactly 3 atoms.

To see how the distributivity operator δ enables us to account for the scopal interaction between universal quantifiers and modified numerals, consider the sentence in (99) below.

(99) Every^x student ate from exactly one^y cake.

It has two possible quantifier scopings, which yield two readings that are truth-conditionally independent:[14]

(*i*) every student x is such that s / he ate from exactly one cake y, possibly different from student to student;

(*ii*) there is exactly one cake y such that every student x ate from it, although it may be that every student ate from more than one cake.

[14] I assume that inverse scope readings are available with modified numerals.

Given the translation schema for universal quantifiers in (93) above, the two possible quantifier scopings of sentence (99) are represented as shown below, and these translations derive the intuitively correct truth conditions. Yet again, the final formulas omit semantically vacuous distributivity operators to improve readability.

(100) a. $\forall x[\text{STUDENT}(x)]\ (\exists^{|y|=1}[\text{CAKE}(y)]\ \delta(\text{EAT-FROM}(x,y)))$

b. $\sigma x(\text{STUDENT}(x)) \wedge$
$\delta x(\sigma y(\text{CAKE}(y) \wedge \delta y(\text{EAT-FROM}(x,y))) \wedge$
$_{|y|=1})$

c. $\sigma x(\text{STUDENT}(x)) \wedge$
$\delta x(\sigma y(\text{CAKE}(y) \wedge \text{EAT-FROM}(x,y)) \wedge$
$|y| = 1)$

(101) a. $\exists^{|y|=1}[\text{CAKE}(y)]\ \delta(\forall x[\text{STUDENT}(x)]\ (\text{EAT-FROM}(x,y)))$

b. $\sigma y(\text{CAKE}(y) \wedge$
$\delta y(\sigma x(\text{STUDENT}(x)) \wedge \delta x(\text{EAT-FROM}(x,y)))) \wedge$
$_{|y|=1}$

c. $\sigma y(\text{CAKE}(y) \wedge$
$\delta y(\sigma x(\text{STUDENT}(x)) \wedge \text{EAT-FROM}(x,y))) \wedge$
$|y| = 1$

The present analysis of cumulative and distributive readings for modified numerals (which can be straightforwardly applied to bare numerals as well) predicts that *ceteris paribus*, cumulative readings are more salient than distributive readings for modified (or bare) numeral sentences. This is because distributive readings require the addition of distributive operators on top of whatever representation we need for cumulative readings.[15] This prediction is confirmed by the experimental results in Gil (1982) and Brooks & Braine (1996); see also Dotlačil (2010) for an overview of these results and critical discussion.[16]

[15] I am indebted to Ede Zimmermann for this observation (p.c.).

[16] As an anonymous *Journal of Semantics* reviewer points out, Brooks & Braine (1996) do not strictly speaking show that sentences with bare numerals are preferably interpreted cumulatively rather than distributively. Rather, they show that such sentences preferably have non-distributive, i.e., collective and branching / cumulative, readings. They test sentences like *Three boys climbed up a tree* in which the branching and cumulative readings collapse. Under the assumption that the branching reading is just a strengthening of the cumulative reading and not an independent reading in its own right, Brooks & Braine (1996) do support the hypothesis that cumulative readings are preferred over distributive readings.

3 Modals and Modified Numerals

This section provides independent evidence for the analysis of modified numerals in terms of post-suppositions. In particular, we will take advantage of post-suppositions and their unusual 'scoping' behavior to account for the unexpected scopal interactions between modified numerals and modal verbs.

To analyze modal verbs, we expand our language and its models in the usual way. We add a set of possible worlds $\mathfrak{W}$ disjoint from $\mathfrak{D}$ and variables over possible worlds w^*, w, w', w_1, w_2 and so on.

In order to give modals an analysis that is parallel to the analysis of quantification over individuals proposed above, we let the domain of possible worlds consists of both atomic worlds and collections / non-atomic worlds, which are simply sets of worlds.

If the reader is not comfortable with this choice because collective readings for modalities do not seem to be available (in contrast to collective readings in the individual domain), s/he should simply regard it as a formal trick whose only purpose is to simplify the logic: if we countenance 'non-atomic worlds', we do not need to introduce distinct variables for worlds on one hand and sets of worlds on the other and distinct types for these two kinds of variables. Everything we will say about modal quantification can be easily reformulated in terms of sets of worlds rather than in terms of 'non-atomic worlds'.

Thus, just as we did for the domain of individuals $\mathfrak{D}$, we will take the domain of possible worlds $\mathfrak{W}$ to be the power set of a given non-empty set WO (this is the set of possible worlds as they are usually conceived of in modal logic): $\mathfrak{W} = \wp^+(\mathsf{WO}) := \wp(\mathsf{WO}) \setminus \{\emptyset\}$.

The sum of two worlds $w \oplus w'$ is the union of the sets w and w'. The part-of relation over worlds $w \leq w'$ is the partial order induced by inclusion $\subseteq$ over the set $\wp^+(\mathsf{WO})$: $w \leq w' := w \subseteq w'$. Atomic worlds are the singleton subsets of WO, identified by means of the predicate $\mathbf{atom}(w) := \forall w' \leq w(w' = w)$. The strict part-of relation is the corresponding order induced by strict inclusion: $w < w' := w \subsetneq w'$.

We relativize the basic interpretation function $\mathfrak{I}$ to *atomic* worlds: for any atomic world $u \in \mathfrak{W}$ and any n-ary relation R, $\mathfrak{I}_u(R) \subseteq \mathfrak{D}^n$.

We take the variable w^* to be the designated variable for the actual world. For any discourse-initial assignment g that we will consider, we assume that $|g(w^*)| = 1$ or equivalently that $\mathbf{atom}(w^*)$ is true. The designated variable w^* performs roughly the same function as the world w in FCS (Heim 1982) style contexts of the form $\langle w, g \rangle$.[17]

Cardinality requirements, random assignment, maximization and

[17] 'Roughly' because, as Philippe Schlenker notes (p.c.), w^* is unshiftable while the implicit world argument used by Heim (1982) is shiftable.

distributivity are defined for possible worlds in the same way that we defined them for individuals:

(102) $|g(w)|$ is the cardinality of the set of atoms in $g(w)$, i.e.: $|g(w)| := |\{u \leq g(w) : \mathbf{atom}(u)\}|$

(103) $[\![|w| = n]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and $|h(w)| = n$

(104) $[\![|w| \leq n]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and $|h(w)| \leq n$

(105) $[\![|w| \geq n]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and $|h(w)| \geq n$

(106) $[\![[w]]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $h[w]g$ and $\zeta = \zeta'$

(107) $[\![\sigma w(\phi)]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff

a. $[\![[w] \wedge \phi]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$

b. there is no h' such that $[\![[w] \wedge \phi]\!]^{\langle g[\zeta],h'[\zeta']\rangle} = \mathbb{T}$ and $h(w) < h'(w)$

(108) $[\![\delta w(\phi)]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and for all atoms $u \leq g(w)$, if we let g' be such that $g'[w]g$ and $g'(w) = u$, then there is a k and a (possibly empty) set of tests $\{\psi_1, \ldots, \psi_m\}$ such that

a. $[\![\phi]\!]^{\langle g'[\zeta],k[\zeta \cup \{\psi_1,\ldots,\psi_m\}]\rangle} = \mathbb{T}$ and

b. $[\![\psi_1 \wedge \ldots \wedge \psi_m]\!]^{\langle k[\zeta],k[\zeta]\rangle} = \mathbb{T}$

We relativize variables / discourse referents for individual $x, y, \ldots$ to possible worlds. That is, in the spirit of Stone (1999), the value of a variable x relative to an assignment g is not an individual, but a partial individual concept, i.e., a partial function from *atomic* worlds to individuals that exist in those atomic worlds. The individuals that are in the range of such partial individual concepts can be atomic or non-atomic.

Thus, for any assignment g and any variable 'over individuals' x, $g(x)$ is in fact not an individual, but a partial function from a non-empty subset (usually a strict subset) of the atomic worlds $\{u \in \mathfrak{W} : \mathbf{atom}(u)\}$ to the domain of individuals $\mathfrak{D}$ (recall that $\mathfrak{D}$ includes both atomic and not-atomic individuals).

Thus, random assignment to variables 'over individuals' is relativized to possible worlds w, as shown below.

(109) $[\![[x_w]]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $h[x]g$, $\zeta = \zeta'$ and $\mathbf{Dom}(h(x)) = \{u \leq h(w) : \mathbf{atom}(u)\}$

Informally, (109) says that introducing a new variable / discourse referent x relative to a world-variable w means that:

- the input assignment g differs from the output assignment h at most with respect to the value of x,

- the input set of post-suppositions ζ is simply passed on to the output context,

- and finally, the partial individual concept assigned to x by the output assignment h is defined only for the atomic worlds u that are part of $h(w)$.

Recall that $h(w) = g(w)$, since $h[x]g$ requires h to differ from g at most with respect to the value assigned to x. Therefore, the last bullet above ultimately says that the partial individual concept assigned to x is defined only for the atomic parts u of the contextually provided world w.

The actual individuals (atomic or not) assigned to each of these atomic worlds u are left completely unconstrained, much as in the extensional system we introduced before.

Maximization / summation over individuals is also relativized to possible worlds w: the partial individual concept $h(x)$ is maximal in a pointwise manner, i.e., relative to each atomic world u in its domain. That is, for any atomic world u that is part of w, the (plural) individual $h(x)(u)$ contains all and only the atomic individuals that satisfy ϕ relative to u.

(110) $[\![\sigma x_w(\phi)]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff

a. $[\![[x_w] \wedge \phi]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$

b. there is no h' such that $[\![[x_w] \wedge \phi]\!]^{\langle g[\zeta],h'[\zeta']\rangle} = \mathbb{T}$ and for any atom $u \leq g(w)$ whatsoever, $h(x)(u) < h'(x)(u)$

Intensional lexical relations are interpreted distributively relative to the world of evaluation w. That is, cumulatively-closed lexical relations are required to hold at every atomic world u that is part of w.

(111) $[\![P_w(x)]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and for each atom $u \leq h(w)$:

- $u \in \mathbf{Dom}(h(x))$
- $h(x)(u) \in {}^*\mathfrak{I}_u(P)$

(112) $[\![R_w(x, y)]\!]^{\langle g[\zeta],h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and for each atom $u \leq h(w)$:

- $u \in \mathbf{Dom}(h(x))$
- $u \in \mathbf{Dom}(h(y))$
- $\langle h(x)(u), h(y)(u)\rangle \in {}^{**}\mathfrak{I}_u(R)$

The definitions of pointwise maximization over individual concepts and cumulatively-closed intensional lexical relations will become clearer if we look at an example. Consider the formula $\sigma x_w(\text{BOOK}_w(x))$.

(113) $[\![\sigma x_w(\text{BOOK}_w(x))]\!]^{\langle g[\zeta], h[\zeta']\rangle} = \mathbb{T}$ iff

- $[\![[x_w] \wedge \text{BOOK}_w(x)]\!]^{\langle g[\zeta], h[\zeta']\rangle} = \mathbb{T}$: we introduce a new partial individual concept $h(x)$ whose domain is the set of atomic worlds $u \leq h(w)$ and require that, relative to each such atomic world u, the (possibly non-atomic) individual $h(x)(u)$ consists of entities that are books in u
  - that is, we introduce a new partial individual concept $h(x)$ such that $\textbf{Dom}(h(x)) = \{u \leq h(w) : \textbf{atom}(u)\}$ and, for all atoms $u \leq h(w)$, $h(x)(u) \in {}^*\mathfrak{I}_u(\text{BOOK})$
  - note that there is an implicit existential commitment (as expected for new discourse-referent introduction): we effectively require each atomic world $u \leq h(w)$ to contain at least some entities that are books
- there is no h' such that $[\![[x_w] \wedge \text{BOOK}_w(x)]\!]^{\langle g[\zeta], h'[\zeta']\rangle} = \mathbb{T}$ and for any atom $u \leq g(w)$ whatsoever, $h(x)(u) < h'(x)(u)$: this is the maximization requirement
  - we previously required $h(x)(u)$ to store *only* books relative to each atomic world $u \leq h(w)$; the maximization requirement makes sure that $h(x)(u)$ stores *all* the books in u, for each atomic world $u \leq h(w)$
  - thus, for each atom $u \leq h(w)$, $h(x)(u)$ is the sum individual containing all and only the books in u and the partial individual concept $h(x)$ is the collection of all these sum individuals relative to their respective atomic worlds

Cardinality requirements for variables 'over individuals' are also relativized to possible worlds, as shown below. For each atomic world u that is part of the world w, the cardinality of x relative to u is n, less than or equal to n, greater than or equal to n etc.

(114) $[\![|x| =_w n]\!]^{\langle g[\zeta], h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and for each atom $u \leq h(w)$:

- $u \in \textbf{Dom}(h(x))$
- $|h(x)(u)| = n$

(115) $[\![|x| \leq_w n]\!]^{\langle g[\zeta], h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and for each atom $u \leq h(w)$:

- $u \in \mathbf{Dom}(h(x))$
- $|h(x)(u)| \leq n$

(116) $[\![|x| \geq_w n]\!]^{\langle g[\zeta], h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and for each atom $u \leq h(w)$:

- $u \in \mathbf{Dom}(h(x))$
- $|h(x)(u)| \geq n$

Finally, proper names are interpreted as rigid designators relative to a possible world w: the partial individual concepts they introduce denote the same individual relative to each atom u in w, formally encoded by the formula below.

(117) $[\![x =_w \text{JASPER}]\!]^{\langle g[\zeta], h[\zeta']\rangle} = \mathbb{T}$ iff $g = h$, $\zeta = \zeta'$ and for each atom $u \leq h(w)$:

- $u \in \mathbf{Dom}(h(x))$
- $h(x)(u)$ is the individual Jasper

As an anonymous reviewer points out, we want negation to discharge post-suppositions in its scope. The simplest examples showing this, e.g., *Peter did not read at least 4 books*, are not very natural but examples of the form *If Peter had not read at least 4 books, ...* are clearly acceptable and interpreted as *If Peter had not read more than 3 books, ...*, i.e., with the cardinality post-supposition trapped within the scope of negation. This can be achieved by defining negation as a distributive modal operator along the lines of Brasoveanu (2010a: p. 497, fn. 24), i.e., by making use of the distributivity operator over worlds defined in (108) above. Since this is not a central concern here, we will not pursue it any further.

3.1 Minimal Requirements

Consider now the sentence in (118) below (based on Nouwen 2010). The most salient reading of this sentence is: the minimum number of books that Jasper is *allowed* to read (if he wants to please his mother) is 10.

(118) Jasper^x should^w read at least ten^y books (to please his$_x$ mother).

As Nouwen (2010) observes, under standard assumptions about the semantics of minimizers and necessity modals, there is no satisfactory analysis of minimal requirements. The reason is that *at least* is analyzed in terms of a minimum operator that retrieves the least number of books Jasper reads in the deontically-ideal worlds quantified over by *should*. If there is at least one deontically-ideal world in which Jasper reads no books, the minimum number is 0. If Jasper reads books in every deontically-ideal world, the minimum number is 1. Either way, the minimum number of books obtained in this way cannot be greater than or equal to 10.

The intuitively correct truth conditions follow automatically in the present framework. The reason is that our analysis of modified numerals is fairly close to their analysis as generalized quantifiers (given the maximization operator they contribute) – and a generalized-quantifier analysis of modified numerals straightforwardly accounts for minimal requirements.[18]

However, analyzing modified numerals as generalized quantifiers fails to account for maximal permissions (to be discussed in the next subsection). This was one of the reasons for their reanalysis as minimizers / maximizers, i.e., in terms of a complex structure involving both a quantifier over degrees and a quantifier over individuals (see Hackl 2000, Heim 2000, Ferreira 2007, Nouwen 2010 and references therein).

That is, the generalized-quantifier analysis of modified numerals captures one type of scopal interactions with modals (minimal requirements), while the minimizer / maximizer analysis captures another type (maximal permissions). We show that the post-suppositional analysis of modified numerals is able to account for both types of interactions.

We analyze necessity modals as distributive universal quantifiers in the modal domain.

(119) $\mathbf{NEC}w(\phi) := \sigma w(R_{w^*}(w)) \wedge \delta w(\phi)$

In (119) above, R is a contextually-provided accessibility relation, i.e., a modal base $+/-$ a built-in ordering source. The formula $R_{w^*}(w)$ is intuitively interpreted as: w is an R-accessible world from the actual world w^*. Just as with lexical relations, we take the formula $R_{w^*}(w)$ to be interpreted in terms of the cumulative closure of the accessibility relation $\mathfrak{I}(R)$. In particular, if u and u' are R-accessible from w^*, then $u \oplus u'$ is also R-accessible from w^*.

Sentence (118) is translated as follows:

(120) a. $\mathbf{NEC}w(\exists x_w[x =_w \text{JASPER}]$
$(\exists^{|y| \geq_w 10}[\text{BOOK}_w(y)]\ (\text{READ}_w(x,y))))$

b. $\sigma w(R_{w^*}(w)) \wedge$
$\delta w([x_w] \wedge x =_w \text{JASPER} \wedge$
$\sigma y_w(\text{BOOK}_w(y) \wedge \text{READ}_w(x,y)) \wedge {}^{|y| \geq_w 10})$

[18] I am indebted to Rick Nouwen for this observation (p.c.).

c. $\sigma w(R_{w^*}(w)) \wedge$
$\delta w([x_w] \wedge x =_w \text{JASPER} \wedge$
$\sigma y_w(\text{BOOK}_w(y) \wedge \text{READ}_w(x,y)) \wedge |y| \geq_w 10)$

The update in (120) instructs us to first introduce all the worlds that are R-accessible from the actual world w^* and store them in w. These are the deontically-ideal worlds that the modal verb *should* universally quantifies over. Then, we distributively check that for each ideal atomic world u in w: if we store Jasper in x and in y all the books that Jasper read in world u, the cardinality of the set of books is at least 10.

That is, Jasper reads at least 10 books in every deontically-ideal world u (since w collects all such deontically-ideal worlds), so we derive the correct truth conditions: the minimum number of books that Jasper is allowed to read (if he wants to please his mother) is 10.

The truth conditions derived by (120) might seem a bit too permissive: they are compatible with a situation in which Jasper should read 20 or more books.[19] I want to suggest that these are, in fact, the correct truth conditions for sentence (118). That is, strictly speaking, sentence (118) is true in a situation in which Jasper should read 20 or more books, but it would be uncooperative for the speaker to use it in such a situation.

That is, I take modified numerals to contribute epistemic implicatures of the kind proposed in Büring (2008) for *at least*. Consider, for example, the sentence in (121) below (from Büring 2008).

(121) Paul has at least four guitars.

This non-modalized sentence is cooperatively used only if the speaker (*i*) is certain that Paul has 4 guitars, (*ii*) considers it possible that Paul has exactly 4 guitars and (*iii*) considers it possible that Paul has more than 4 guitars.

Similarly, the modalized sentence in (118) is cooperatively used only if the speaker considers it possible that reading exactly 10 books is enough for Jasper to please his mother.

Consider also the example below from Nouwen (2010) (see also Geurts & Nouwen 2007 and Krifka 2007):

(122) Jasper invited maximally fifty people to his party.

Following the proposal in Büring (2008), this sentence is cooperatively used only if the speaker (*i*) is certain that Jasper did not invite more than 50 people to his party, (*ii*) considers it possible that Jasper invited exactly 50 people to his party and (*iii*) considers it possible that Jasper

[19] I am indebted to Rick Nouwen for this observation (p.c.).

invited less than 50 people to his party. That is, if (122) is cooperatively uttered, it indicates that the speaker does not know exactly how many people Jasper invited. So, (122) is interpreted as being about the non-trivial (i.e., non-singleton) *range* of cardinalities possible at that point in discourse. It is therefore unacceptable for a speaker to utter (122) and continue with: *43, to be precise*.[20]

Thus, it seems that we do not need to derive the modal flavor of this kind of non-modalized indicative sentences by inserting covert modals or covert speech act operators (as Geurts & Nouwen 2007, Krifka 2007 and Nouwen 2010 suggest).

3.2 Maximal Permissions

This subsection shows that we correctly analyze maximal permissions, e.g., (123) below (from Nouwen 2010). The most salient reading of this sentence is: the maximum number of people Jasper is allowed to invite is 10.

(123) Jasper^x is allowed^w to invite at most ten^y people.

We take possibility modals to be the modal counterpart of maximal *some*. Furthermore, we analyze *some* and *might* / *allow* etc. in a way that is parallel to modified numerals, i.e., in terms of a maximization operator σ followed by a post-suppositional cardinality requirement.

The maximization operator σ is justified by the maximal (E-type) anaphora exemplified by the following well-known examples (see Evans 1977, 1980 and Roberts 1987, 1989).

(124) Harry bought some sheep. Bill vaccinated them.

(125) A wolf might come in. It would eat Jasper first.

The most salient reading of (124) is that Bill vaccinated *all* the sheep that Harry bought. Similarly, the most salient reading of (125) is that for *every* epistemically-possible scenario of a wolf coming in, the wolf eats Jasper first.

The translation schemas for *some* and possibility modals are provided below.

(126) *some* (extensional version) $\exists^{|x|\geq 1}[\phi]\ (\psi) := \sigma x(\phi \wedge \psi) \wedge\ ^{|x|\geq 1}$

(127) $\mathbf{POS}w(\phi) := \exists^{|w|\geq 1}[R_{w^*}(w)]\ (\phi)$
$:= \sigma w(R_{w^*}(w) \wedge \phi) \wedge\ ^{|w|\geq 1}$

[20] As Rick Nouwen points out (p.c.), it seems that only class B modified numerals trigger such modalized implicatures, but not class A modified numerals. A discourse like *Jasper invited fewer than fifty people to his party. 43, to be precise.* is felicitous.

Sentence (123) is translated as follows:

(128) a. $\mathbf{POS}w(\exists x_w[x =_w \text{JASPER}]$
$(\exists^{|y| \leq_w 10}[\text{PERSON}_w(y)]\ (\text{INVITE}_w(x, y))))$

b. $\sigma w(R_{w^*}(w) \wedge [x_w] \wedge x =_w \text{JASPER} \wedge$
$\sigma y_w(\text{PERSON}_w(y) \wedge \text{INVITE}_w(x, y)) \wedge {}^{|y| \leq_w 10}) \wedge$
${}^{|w| \geq 1}$

c. $\sigma w(R_{w^*}(w) \wedge [x_w] \wedge x =_w \text{JASPER} \wedge$
$\sigma y_w(\text{PERSON}_w(y) \wedge \text{INVITE}_w(x, y))) \wedge$
${}^{|y| \leq_w 10} \wedge {}^{|w| \geq 1}$

d. $\sigma w(R_{w^*}(w) \wedge [x_w] \wedge x =_w \text{JASPER} \wedge$
$\sigma y_w(\text{PERSON}_w(y) \wedge \text{INVITE}_w(x, y))) \wedge$
$|y| \leq_w 10 \wedge |w| \geq 1$

The update in (128) instructs us to introduce all the worlds w that are R-accessible from the actual world w^* (i.e., deontically-ideal) such that Jasper invites some people in w. For each world w, we store in y all the people invited by Jasper. Finally, we check that there is at least 1 such ideal world w and that the cardinality of the set y of invited people in *each* world w taken individually is at most 10.

The 'scoping' behavior of post-suppositions is crucial for the derivation of the correct truth conditions. This is what enables us to go from the formula in (128b) above to the formula in (128c), where the cardinality requirement ${}^{|y| \leq_w 10}$ contributed by the narrow-scope modified numeral is 'scoped out' from underneath the possibility modal. This is parallel to the way in which we derive cumulative readings with non-increasing modified numerals.

Once the cardinality requirement ${}^{|y| \leq_w 10}$ is 'scoped out', we can substitute the post-suppositions with at-issue tests *salva veritate*, as shown in (128d).

3.3 Distributive Permissions

Analyzing possibility modals in parallel to modified numerals predicts that they can also have distributive readings of the following form:

(129) $\mathbf{POS}w(\delta(\phi)) := \exists^{|w| \geq 1}[R_{w^*}(w)]\ \delta(\phi)$
$:= \sigma w(R_{w^*}(w) \wedge \delta w(\phi)) \wedge {}^{|w| \geq 1}$

The resulting distributive translation of sentence (123) is given below.

(130) a. $\mathbf{POS}w(\delta(\exists x_w[x =_w \text{JASPER}]$
$(\exists^{|y| \leq_w 10}[\text{PERSON}_w(y)]\ (\text{INVITE}_w(x, y)))))$

b. $\sigma w(R_{w^*}(w) \wedge$
$\quad \delta w([x_w] \wedge x =_w \textsc{jasper} \wedge$
$\quad\quad \sigma y_w(\textsc{person}_w(y) \wedge \textsc{invite}_w(x, y)) \wedge {}^{|y| \leq_w 10})) \wedge$
$\quad {}_{|w| \geq 1}$

c. $\sigma w(R_{w^*}(w) \wedge$
$\quad \delta w([x_w] \wedge x =_w \textsc{jasper} \wedge$
$\quad\quad \sigma y_w(\textsc{person}_w(y) \wedge \textsc{invite}_w(x, y)) \wedge |y| \leq_w 10)) \wedge$
$\quad |w| \geq 1$

The update in (130) is interpreted as: there is at least 1 world w that is R-accessible from the actual world w^* such that the maximum number of people Jasper invites in w is at most 10. That is, inviting at most 10 people is something that Jasper is allowed to do. This is a consequence of the fact that the distributivity operator δ over the nuclear scope of the possibility modal forces the post-suppositions to be interpreted '*in situ*'.

This rather weak reading is not intuitively available for sentence (123). We will follow Nouwen (2010) and assume that such readings are blocked by the availability of (and competition with) the parallel construction with a bare numeral instead of a modified numeral.

In general, the proposal that bare numerals can block modified numerals predicts that whenever (*i*) an operator, e.g., **POS***w*, can have both a cumulative and a distributive reading and (*ii*) this operator has a modified numeral in its scope, the distributive reading that locally discharges the post-supposition contributed by the modified numeral competes with and is blocked by the parallel construction with a bare numeral. The rest of this subsection attempts to sketch how this blocking mechanism is supposed to work; many details and various related issues and examples are not addressed by this preliminary sketch.[21]

The reason is that the post-supposition contributed by the modified numeral is trapped '*in situ*' and behaves like it is part of the regular at-issue meaning, so it is more economical (in a sense that will remain unspecified here) to instead use the bare numeral, which has no post-suppositional component to begin with.

For concreteness, the bare numeral counterpart of sentence (123) is provided in (131) below. Just as before, the possibility modal can be interpreted cumulatively or distributively, as shown in (132) and (133) respectively. Either way, we obtain the same reading: there is at least 1 world w that is R-accessible from the actual world w^* such that Jasper invites 10 people in w. That is, inviting 10 people is something that Jasper is allowed to do.

[21] I am indebted to an anonymous reviewer for emphasizing this point and for her / his insightful comments and suggestions about this subsection.

(131) Jasper^x is allowed^w to invite ten^y people.

(132) a. $\mathbf{POS}w(\exists x_w[x =_w \text{JASPER}]$
$(\exists y_w[|y| =_w 10 \wedge \text{PERSON}_w(y)]$
$(\text{INVITE}_w(x, y))))$

b. $\sigma w(R_{w^*}(w) \wedge [x_w] \wedge x =_w \text{JASPER} \wedge$
$[y_w] \wedge |y| =_w 10 \wedge \text{PERSON}_w(y) \wedge \text{INVITE}_w(x, y))$
$\wedge |w| \geq 1$

(133) a. $\mathbf{POS}w(\delta(\exists x_w[x =_w \text{JASPER}]$
$(\exists y_w[|y| =_w 10 \wedge \text{PERSON}_w(y)]$
$(\text{INVITE}_w(x, y)))))$

b. $\sigma w(R_{w^*}(w) \wedge$
$\delta w([x_w] \wedge x =_w \text{JASPER} \wedge$
$[y_w] \wedge |y| =_w 10 \wedge \text{PERSON}_w(y) \wedge \text{INVITE}_w(x, y)))$
$\wedge |w| \geq 1$

Importantly, blocking goes through if we take into account the epistemic implicatures associated with sentence (123) and not only its at-issue content.[22] As far as the at-issue content is concerned, (132) and (133) say that inviting 10 people is something that Jasper is allowed to do. In contrast, (130) says that inviting 10 people or less is something that Jasper is allowed to do.

Thus, (132) and (133) are false if Jasper invites exactly 5 people in every *R*-accessible world (that is, he may and must invite exactly 5 people), while (130) is true in this case – *as far as its at-issue content is concerned*. However, sentence (123) (under the reading in (130)) would not be used cooperatively in this case (see the discussion of Büring 2008 and epistemic implicatures above) since the modified numeral *at most ten^y people* indicates that the speaker considers it possible that Jasper is allowed to invite exactly 10 people. Covertly enriching the meaning of modified numerals in this way, i.e., by taking into account epistemic implicatures, is similar to the strategy employed in Nouwen (2010), where covert ambiguity is the driving force behind some of the cases in which blocking applies.

As an anonymous reviewer points out, even when (123) is enriched with epistemic implicatures, it still does not say the exact same thing as (131): (123) says that for some number $n \leq 10$, Peter is allowed to invite n people – and both $n = 10$ and $n < 10$ are live possibilities; in contrast, (131) says something about $n = 10$ but it does not say anything about numbers $n < 10$. Given that issues related to the exact form of the blocking mechanism are sufficiently far from the core proposal of this paper, we will leave this problem open here.

[22] I am indebted to an anonymous reviewer and Rick Nouwen (p.c.) for insightful discussions of this issue.

Assuming we can ultimately provide a satisfactory characterization of the blocking mechanism, we could also use it to derive the infelicity of minimal permissions like sentence (134) below: the distributive reading of the possibility modal in (134) is blocked by the bare numeral construction in (135).

(134) #A course is allowed to have at least four registered students (to be approved by the administration).

(135) A course is allowed to have four registered students.

(136) A course must have at least four registered students (to be approved by the administration).

The cumulative reading of the possibility modal in (134) is presumably blocked by the alternative, unambiguous construction in (136), where the possibility modal is replaced with its unambiguously distributive universal counterpart.

Interestingly, the cumulative reading of the possibility sentence in (134) is blocked by the necessity sentence in (136), but the cumulative reading of (123) above seems to not be blocked by its necessity counterpart in (137) below. Once again, we will leave this issue open.

(137) Jasper^{x} is required^{w} to invite at most ten^{y} people.

In sum, we have shown in this section – see subsections 3.1 and 3.2 in particular – that an analysis of modified numerals in terms of quantificational alternatives and post-suppositions enables us to correctly account for the puzzling scopal interactions between modified numerals and modals noticed in the previous literature (see Krifka 1999 for a related discussion of interactions with attitude reporting verbs).

This account of the scopal interactions between modals and modified numerals provides an independent justification for the analysis of modified numerals in terms of post-suppositions. More broadly, in view of the systematic patterns of interaction between modified numerals and other quantificational expressions, this section provides additional justification for analyzing modified numerals in terms of quantificational and not focus alternatives.

4 Cumulative Readings for Universal Quantifiers

This final contentful section outlines a way in which the present account of cumulativity can be generalized to capture the fact that distributive universal quantifiers can also have cumulative readings, as observed in Schein (1993), Kratzer (2000) and Champollion (2010) among others.

Consider the sentence below:

(138) Three^x copy editors (between them) caught every^y mistake in the manuscript.

Sentence (138) is cumulatively interpreted as: there are three copy editors such that each of them caught at least one mistake and every mistake was caught by at least one of the three editors.

Given the analysis of distributive universal quantifiers proposed above, sentence (138) is translated as shown in (139) below. For expository simplicity, we revert to the extensional version of the system without post-suppositions throughout this section.

(139) a. $\exists x[|x| = 3 \wedge \text{EDITOR}(x)]\ (\forall y[\text{MISTAKE}(y)]\ (\text{CATCH}(x, y)))$

b. $[x] \wedge |x| = 3 \wedge \text{EDITOR}(x) \wedge$
$\sigma y(\text{MISTAKE}(y)) \wedge \delta y(\text{CATCH}(x, y))$

However, the update in (139) does not derive the intuitively correct cumulative reading, but the distributive reading of sentence (138): each mistake y is such that (each of) the three editors x caught it. The representation that captures the cumulative reading does not have a distributivity operator δy, as shown in (140) below.

(140) $[x] \wedge |x| = 3 \wedge \text{EDITOR}(x) \wedge$
$\sigma y(\text{MISTAKE}(y)) \wedge \text{CATCH}(x, y)$

One way out would be to say that universal quantifiers of the form *every+NP* are only optionally distributive, i.e., only optionally contribute a distributivity operator δy over their nuclear scope, just as we took modified numerals to be only optionally distributive in subsection 2.4 above. However, this would make the incorrect prediction that *every+NP* quantifiers can have collective readings much like modified numerals can. And although *everyone* and modified numerals can have collective readings, as shown in (141) and (142) below, *every+NP* cannot.[23]

(141) Everyone gathered in the park.

[23] A COCA search for *every* followed by *gather* within a 4-word window revealed about 20 examples, but none of them had the required form, i.e., the same syntactic structure as the intuitively infelicitous example in (143) above. In contrast, there are about 20 COCA examples of the required form with *everyone* followed by *gather* within a 4-word window – and 2 examples of the form *At least n students gather*, resulting from a search for *at least* + CARDINAL-NUMERAL followed by *gather* within a 4-word window.

(142) Exactly / At least three students gathered in the park.

(143) #Every student gathered in the park.

The optional-distributivity account of cumulative readings for universal quantifiers also makes incorrect predictions with respect to other examples that do not involve collective readings. Consider the example below from Kratzer (2000).

(144) Every^x copy editor caught 500^y mistakes in the manuscript.

As Kratzer (2000) notes, this sentence does not have a cumulative reading to the effect that between them, the copy editors caught a total of 500 mistakes in the manuscript. The only available reading is the distributive one: every copy editor is such that s / he caught 500 mistakes. However, the optional-distributivity account incorrectly predicts that a cumulative reading is in fact available for this example.

Finally, the optional-distributivity account also fails to generalize to the mixed cumulative-distributive sentence below from Schein (1993):

(145) Three^x video games taught every^y quarterback two^z new plays.

As Kratzer (2000) observes, *every^y quarterback* and *three^x video games* are related cumulatively: between them, a total of three video games taught all the quarterbacks. But *every^y quarterback* behaves just like an ordinary distributive quantifier with respect to *two^z new plays*: every quarterback learned two possibly different plays. The optional-distributivity account can only derive an across-the-board distributive or an across-the-board cumulative reading for this sentence, but not the correct mixed (cumulative-distributive) reading.

4.1 Evaluation Pluralities vs Ontological Pluralities

We will therefore outline an alternative way to incorporate pluralities into a dynamic system and capture the fact that universal quantifiers can have cumulative and mixed cumulative-distributive readings.

The basic idea is that, instead of enriching our ontology with plural / non-atomic individuals, we can enrich our contexts of evaluation and take them to consist of sets of variable assignments instead of single assignments (see van den Berg 1996, Nouwen 2003, Wang 2005 and Brasoveanu 2007 among others for more discussion). That is, instead of adding ontological / domain-level pluralities, we add evaluation / discourse-reference-level pluralities.

The models we will work with are exactly like the FOL models, i.e., $\mathfrak{M} = \langle \mathfrak{D}, \mathfrak{I} \rangle$, where $\mathfrak{D}$ is the domain of individuals and $\mathfrak{I}$ is the basic interpretation function such that $\mathfrak{I}(R) \subseteq \mathfrak{D}^n$ for any n-ary relation R.

Importantly, the domain of individuals $\mathfrak{D}$ consists of atomic individuals *only*.

An $\mathfrak{M}$-assignment g is a total function from the set of variables $\mathcal{V}$ to $\mathfrak{D}$. As we already remarked, the essence of quantification in FOL is pointwise / variablewise manipulation of variable assignments, abbreviated $h[x]g$: h differs from g at most with respect to the value it assigns to x. We generalize this to sets of assignments $H[x]G$ cumulative-quantification style, as shown below.

(146) $$H[x]G := \begin{cases} \text{for all } h \in H, \text{ there is a } g \in G \text{ such that } h[x]g \\ \text{for all } g \in G, \text{ there is a } h \in H \text{ such that } h[x]g \end{cases}$$

Formally, this is a natural generalization: $H[x]G$ is an equivalence relation over sets of assignments, just as $h[x]g$ is an equivalence relation over single assignments.

A set of assignments G can be represented as a matrix. The rows of the matrix represent variable assignments g_1, g_2, g_3 etc. The columns represent variables x, y etc. The objects in the cells of the matrix are values that assignments assign to variables: $boy_1 = g_1(x)$, $boy_2 = g_2(x)$, $movie_1 = g_1(y)$, $movie_2 = g_2(y)$ etc.

(147)

G	...	x	y	...
g_1	...	boy_1	$movie_1$	...
g_2	...	boy_2	$movie_2$	...
g_3	...	boy_3	$movie_3$	...
...	...	...	...	...

or simply:

...	x	y	...
...	boy_1	$movie_1$	...
...	boy_2	$movie_2$	...
...	boy_3	$movie_3$	...
...	...	...	...

Just as in DPL, formulas denote binary relations between input and output contexts. But these contexts are now sets of assignments instead of single assignments.

Lexical relations are tests – again, just as in DPL. They require the output context H to be the same as the input context G (i.e., they simply pass on the input context) and check that H satisfies the lexical relation R in a distributive way. That is, each assignment $h \in H$ is required to satisfy R. Note that we do not make any use of the cumulative-closure operators * / **.

(148) $[\![R(x_1,\ldots,x_n)]\!]^{\langle G,H\rangle} = \mathbb{T}$ iff $G = H$ and for all $h \in H$, $\langle h(x_1),\ldots,h(x_n)\rangle \in \mathfrak{I}(R)$

A typical output set of assignments $H = \{h, h', h'', \ldots\}$ that satisfies the atomic formula $R(x_1,\ldots,x_n)$ in such a distributive way is provided in (149) below.

(149)

H	...	x_1	...	x_n	...
h	...	$\alpha_1\ (= h(x_1))$	...	$\alpha_n\ (= h(x_n))$	...
		$\langle\alpha_1,\ldots,\alpha_n\rangle \in \mathfrak{I}(R)$			
h'	...	$\beta_1\ (= h'(x_1))$	...	$\beta_n\ (= h'(x_n))$	...
		$\langle\beta_1,\ldots,\beta_n\rangle \in \mathfrak{I}(R)$			
h''	...	$\gamma_1\ (= h''(x_1))$	...	$\gamma_n\ (= h''(x_n))$	...
		$\langle\gamma_1,\ldots,\gamma_n\rangle \in \mathfrak{I}(R)$			
...	...	...	...	...	...

Just as before, cardinality constraints on the values of a variable x are tests, as shown by the semantic clauses below. But now we collect all the values that the current context of evaluation stores in column x (since there are no plural individuals!) and then place requirements on the cardinality of this set of values. The set that is the result of collecting all the values stored in column x of a matrix G is abbreviated as $G(x)$.

(150) $G(x) := \{g(x) : g \in G\}$

(151) $|G(x)|$ is the cardinality of the set of individuals $G(x)$

(152) $[\![|x| = n]\!]^{\langle G,H\rangle} = \mathbb{T}$ iff $G = H$ and $|H(x)| = n$

(153) $[\![|x| \leq n]\!]^{\langle G,H\rangle} = \mathbb{T}$ iff $G = H$ and $|H(x)| \leq n$

(154) $[\![|x| \geq n]\!]^{\langle G,H\rangle} = \mathbb{T}$ iff $G = H$ and $|H(x)| \geq n$

Dynamic conjunction and random assignment are defined DRT / FCS / DPL style. In particular, dynamic conjunction is interpreted as relation composition.

(155) $[\![\phi \wedge \psi]\!]^{\langle G,H\rangle} = \mathbb{T}$ iff there is a K such that $[\![\phi]\!]^{\langle G,K\rangle} = \mathbb{T}$ and $[\![\psi]\!]^{\langle K,H\rangle} = \mathbb{T}$

(156) Random assignment:
$[\![[x]]\!]^{\langle G,H\rangle} = \mathbb{T}$ iff $H[x]G$

4.2 Bare Numerals and Singular Indefinites, Evaluation-Plurality Style

Unsurprisingly, the format for the translation of singular indefinite articles and bare numerals remains the same as before. This is because the only real change to the dynamic system is the way we incorporate pluralities in our logic.

(157) $\exists x[|x| = n \wedge \phi]\ (\psi)$

For singular indefinite articles, n is 1. For the bare numeral *two*, n is 2. For the bare numeral *three*, n is 3 etc. Two example translations are provided below – and they are identical to the ones in (26) and (27) above.

(158) A^x wolf came in. $\rightsquigarrow \exists x[|x| = 1 \wedge \text{WOLF}(x)]\ (\text{COME-IN}(x))$

(159) Two^x wolves came in. $\rightsquigarrow \exists x[|x| = 2 \wedge \text{WOLF}(x)]\ (\text{COME-IN}(x))$

Just as before, we 'decompose' the translation schema for bare numerals and singular indefinites into the same flat conjunction of elementary formulas. The only difference is that these elementary formulas are interpreted according to the new evaluation-plurality based semantic clauses provided above.

(160) $\exists x[|x| = n \wedge \phi]\ (\psi) := [x] \wedge |x| = n \wedge \phi \wedge \psi$

Proper names also receive the same translation.

(161) $\exists x[x = \text{JASPER}]\ (\phi) := [x] \wedge x = \text{JASPER} \wedge \phi$

Just as before, pronouns are indexed with the variable introduced by their antecedent and their translation is that variable itself.

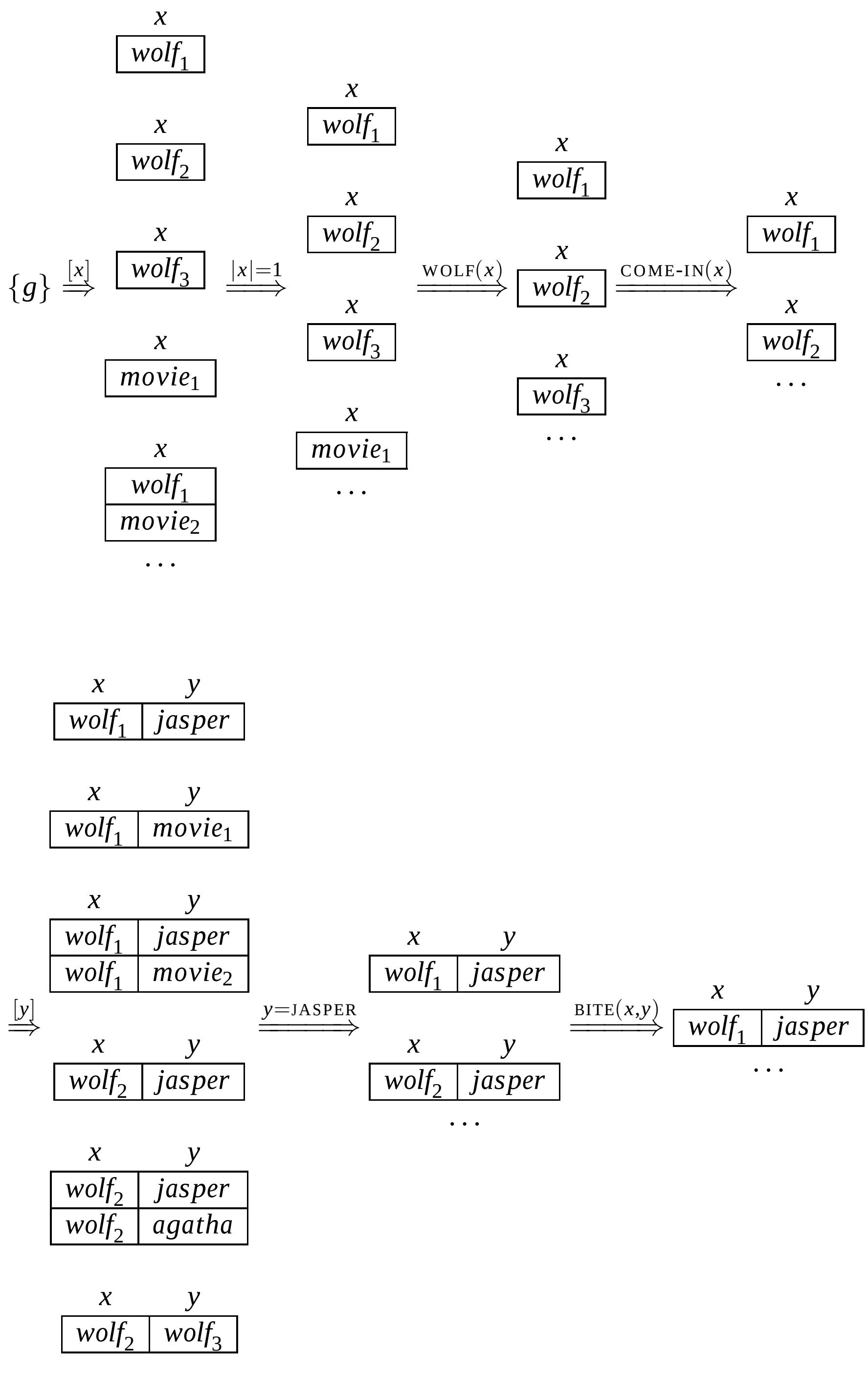
The two-sentence discourse in (30) above, repeated below for convenience, is also compositionally translated just as we translated it before.

(162) A^x wolf came in. It_x bit Jasper^y.

(163) a. $\exists x[|x| = 1 \wedge \text{WOLF}(x)]\ (\text{COME-IN}(x)) \wedge$
$\exists y[y = \text{JASPER}]\ (\text{BITE}(x, y))$

b. $[x] \wedge |x| = 1 \wedge \text{WOLF}(x) \wedge \text{COME-IN}(x) \wedge$
$[y] \wedge y = \text{JASPER} \wedge \text{BITE}(x, y)$

The actual interpretation of the formulas is different, however. Suppose, for simplicity, that our input context G is the singleton set $\{g\}$, where g assigns some arbitrary values to all variables. The conjunction of formulas in (163b) above updates this input context as shown in (164) below. It is instructive to compare the sequence of evaluation-plurality based updates depicted in (164) and the corresponding ontological-plurality based updates depicted in (32) above.

(164) 

The update in (164) proceeds as follows. We first introduce x, i.e., assign it a random value. The result: many matrices, some containing only one row, some containing two rows etc. and assigning all possible individuals or combinations thereof to x. That is, we now have a graph with many paths. Then, the test $|x| = 1$ eliminates some of the paths in the graph, namely all those paths that end in a matrix assigning more than one entity to x. That is, any matrix with multiple rows is eliminated). The test $\text{WOLF}(x)$ eliminates further paths in the graph, namely all those that end in a matrix where x is not assigned a wolf. The test $\text{COME-IN}(x)$ eliminates all the wolves that didn't come in.

Then, we introduce another variable y that extends the graph in many different ways. The subsequent test $y = \text{JASPER}$ prunes down the graph by eliminating all the matrices that don't assign Jasper to y. Finally, the test $\text{BITE}(x, y)$ keeps only the matrices H such that, for any row $h \in H$, the individual $h(x)$ bit the individual $h(y)$.

Except for the fact that we allow matrices with multiple rows, the interpretation graph in (164) is not different from the one in (32) – and is not different from the way interpretation proceeds in classical FOL or classical DRT / FCS.

Just as before, we can depict updates by choosing a single, typical path through the graph:

(165) $\{g\} \xrightarrow{[x]\wedge|x|=1\wedge\text{WOLF}(x)\wedge\text{COME-IN}(x)} \begin{array}{c} x \\ \hline \textit{wolf}_1 \end{array}$

$\xrightarrow{[y]\wedge y=\text{JASPER}\wedge\text{BITE}(x,y)} \begin{array}{c|c} x & y \\ \hline \textit{wolf}_1 & \textit{jasper} \end{array}$

The definition of truth is also the same as before: it says that a formula is true if there is at least one successful path through the graph / binary relation denoted by ϕ.

(166) Truth: a formula ϕ is true relative to an input set of assignments G iff there is an output set of assignments H such that $[\![\phi]\!]^{\langle G,H\rangle} = \mathbb{T}$.

We get cumulative readings for bare numerals automatically, as shown below. Note that the final, rightmost output context in (39) is just the matrix representation of part of Figure 1: each assignment / row in this matrix corresponds to one of the 'seeing'-arrows in Figure 1.

(167) Three^x boys saw five^y movies.

(168) a. $\exists x[|x| = 3 \wedge \text{BOY}(x)]\ (\exists y[|y| = 5 \wedge \text{MOVIE}(y)]\ (\text{SEE}(x, y)))$

b. $[x] \wedge |x| = 3 \wedge \text{BOY}(x) \wedge [y] \wedge |y| = 5 \wedge \text{MOVIE}(y) \wedge \text{SEE}(x, y)$

(169) $\{g\} \xrightarrow{[x]\wedge|x|=3\wedge\text{BOY}(x)} \begin{array}{c} x \\ \hline \textit{boy}_1 \\ \textit{boy}_3 \\ \textit{boy}_4 \end{array}$

$$\xrightarrow{[y]\wedge|y|=5\wedge\text{MOVIE}(y)\wedge\text{SEE}(x,y)}$$

x	y
boy_1	$movie_1$
boy_3	$movie_2$
boy_3	$movie_3$
boy_3	$movie_4$
boy_4	$movie_4$
boy_4	$movie_5$

Importantly, the lexical relations $\text{BOY}(x)$, $\text{MOVIE}(y)$ and $\text{SEE}(x,y)$ are not cumulatively closed: there are no * or ** operators. Lexical relations are distributively interpreted relative to their input set of assignments, i.e., they relate atomic individuals as in classical FOL. We can do this and still capture cumulative readings for bare numerals because we cumulate in the meta-language, at the level of contexts of evaluation.

4.3 Universal Quantifiers

We are now able to translate distributive universal quantification as in (170) below. A universal quantifier introduces the set of all individuals x that satisfy the restrictor ϕ – by means of $\mathbf{M}x(\phi)$ – and then checks that *each* of these individuals also satisfies the nuclear scope ψ – by means of $\mathbf{D}x(\psi)$.

(170) $\forall x[\phi]\ (\psi) := \mathbf{M}x(\phi) \wedge \mathbf{D}x(\psi)$

The maximization operator $\mathbf{M}x$ is the evaluation-plurality counterpart of the Link-style σx operator we introduced for ontological pluralities. Similarly, the distributivity operator $\mathbf{D}x$ is the counterpart of the δx operator.[24]

(171) $[\![\mathbf{M}x(\phi)]\!]^{\langle G,H\rangle} = \mathbb{T}$ iff

a. $[\![[x] \wedge \phi]\!]^{\langle G,H\rangle} = \mathbb{T}$

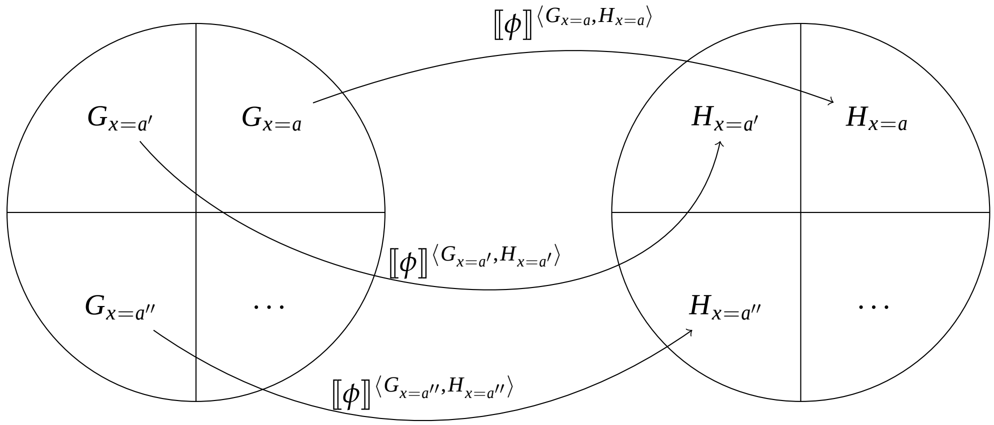
b. there is no H' such that $H(x) \subsetneq H'(x)$ and $[\![[x] \wedge \phi]\!]^{\langle G,H'\rangle} = \mathbb{T}$

(172) $G_{x=a} := \{g \in G : g(x) = a\}$

(173) $[\![\mathbf{D}x(\phi)]\!]^{\langle G,H\rangle} = \mathbb{T}$ iff $G(x) = H(x)$ and for any $a \in G(x)$, $[\![\phi]\!]^{\langle G_{x=a},H_{x=a}\rangle} = \mathbb{T}$

(174) Updating the set of assignments G with a formula ϕ distributively over x:

[24] I am grateful to an anonymous *Journal of Semantics* reviewer for a very helpful and detailed discussion of the relative merits of the unselective distributivity operator I had in a previous version of the paper (and in Brasoveanu 2010b) and the selective distributivity operator $\mathbf{D}x$ that is used here.



Definition (173) states that updating a set of assignments G with a formula ϕ distributively over a variable x means:

(*i*) generating the x-partition of G, i.e., $\{G_{x=a} : a \in G(x)\}$,

(*ii*) updating each cell $G_{x=a}$ in the partition with the formula ϕ,

(*iii*) and finally, taking the union of the resulting output sets of assignments.

The first conjunct $G(x) = H(x)$ in (173) is required to ensure that there is a bijection between the partition induced by the variable x over the input context G and the one induced over the output context H. Without this requirement, we could introduce arbitrary new values for x in the output context H, i.e., arbitrary new partition cells.[25] The second conjunct is the one that actually defines the distributive update: the formula ϕ relates every partition cell in the input context G to the corresponding partition cell in the output context H.

We can now capture the cumulative reading of sentence (138) above. This sentence is translated as shown in (175) below.

(175) a. $\exists x[|x| = 3 \wedge \text{EDITOR}(x)]\ (\forall y[\text{MISTAKE}(y)]\ (\text{CATCH}(x, y)))$

b. $[x] \wedge |x| = 3 \wedge \text{EDITOR}(x) \wedge$
$\mathbf{M}y(\text{MISTAKE}(y)) \wedge \mathbf{D}y(\text{CATCH}(x, y))$

c. $[x] \wedge |x| = 3 \wedge \text{EDITOR}(x) \wedge$
$\mathbf{M}y(\text{MISTAKE}(y)) \wedge \text{CATCH}(x, y)$

We introduce a set x of three editors and the set y of all mistakes and check that for every assignment h in the resulting output context H, the editor $h(x)$ caught the mistake $h(y)$.

[25] Nouwen (2003) was the first to observe that we need to add the first conjunct in (173) to the original definition of distributivity in van den Berg (1996).

Crucially, letting universal quantifiers distribute over evaluation pluralities, as opposed to ontological pluralities, allows them to have cumulative readings (and does not incorrectly predict that they can have collective readings). The reason is that evaluation-level distributivity distributes over variable assignments at the same time as it distributes over the values assigned to the targeted variable. Therefore, we operate over both sets of entities (editors and mistakes) that we need to cumulate over simultaneously. In fact, the distributivity operator $\mathbf{D}y$ is semantically vacuous in (175a) / (175b) and it can be omitted – as we did in (175c).

The distributivity operator **D** is not always vacuous, however. Consider again the example from Kratzer (2000) in (144) above. As we already observed, this sentence does not have a cumulative reading to the effect that between them, the copy editors caught a total of 500 mistakes in the manuscript. The only available reading is the distributive one: every copy editor is such that s / he caught 500 mistakes.

We derive the distributive reading if the universal quantifier takes scope over the numeral. That is, the evaluation-plurality based analysis of universal quantifiers automatically restricts the availability of cumulative readings: they are possible only if universal quantifiers have narrow scope relative to the numerals they 'cumulate' with. As long as the non-surface scope *500>>every* is somehow blocked for sentence (144), we correctly derive the unavailability of the cumulative reading.

The translation of the surface-scope reading *every>>500* for sentence (144) is provided in (176) below.

(176) a. $\forall x[\text{EDITOR}(x)]\ (\exists y[|y| = 500 \wedge \text{MISTAKE}(y)]\ (\text{CATCH}(x, y)))$

b. $\mathbf{M}x(\text{EDITOR}(x)) \wedge$
$\mathbf{D}x([y] \wedge |y| = 500 \wedge \text{MISTAKE}(y) \wedge \text{CATCH}(x, y))$

We introduce the set of all copy editors x and we check that *each* of them caught 500 mistakes. The distributivity operator **D** is not semantically vacuous in this case, so it cannot be omitted.

Finally, the analysis of universal quantifiers as distributors over evaluation pluralities also generalizes to the mixed cumulative-distributive sentence in (145) above. We capture the correct reading if we preserve the surface-scope relations between the three quantifiers: *three* >> *every* >> *two*. The resulting translation, which derives the intuitively correct truth conditions, is provided below.

(177) a. $\exists x[|x| = 3 \wedge \text{GAME}(x)]$
$(\forall y[\text{Q.BACK}(y)]$
$(\exists z[|z| = 2 \wedge \text{PLAY}(z)]\ (\text{TEACH}(x, y, z))))$

b. $[x] \wedge |x| = 3 \wedge \text{GAME}(x) \wedge \mathbf{M}y(\text{Q.BACK}(y)) \wedge$
$\mathbf{D}y([z] \wedge |z| = 2 \wedge \text{PLAY}(z) \wedge \text{TEACH}(x, y, z))$

We will not further pursue this evaluation-plurality dynamic system here or try to integrate the two kinds of pluralities into a single system (see Brasoveanu 2008 for a way to do that). A system countenancing both evaluation pluralities (sets of assignments) and ontological pluralities (non-atomic individuals) is needed[26] to capture the meaning of universal quantifiers like *every*^{x} *three houses* among other things. The output matrix resulting after the interpretation of such a quantifier would store a non-atomic individual consisting of three houses in each x-cell and would have as many cells as there are non-atomic individuals that have three houses as their atomic parts.

We only mention that post-suppositions and the resulting analysis of cumulative readings for modified numerals proposed above, together with the account of the scopal interactions between modals and modified numerals, can be straightforwardly incorporated into a dynamic system based on evaluation pluralities instead of ontological pluralities. A brief sketch of how this can be done is provided in Brasoveanu (2010b).

In sum, we ended up distinguishing between two kinds of pluralities in this section, namely (*i*) evaluation pluralities, i.e., sets of assignments, and (*ii*) ontological pluralities, i.e., non-atomic individuals (which we would ultimately want to allow alongside evaluation pluralities). The maximization operator $\mathbf{M}x$ and the distributivity operator $\mathbf{D}x$, needed for distributive quantification, are to evaluation pluralities what the familiar Link-style sum and distributivity operators σx and δx are to ontological pluralities. The bigger picture that seems to emerge is that cumulativity is just non-distributivity with respect to evaluation pluralities, while collectivity (group readings, 'partial covers' etc.) is just non-distributivity with respect to domain pluralities.

5 Conclusion

The goal of the paper was to provide an account of cumulative readings with non-increasing modified numerals. To this effect we introduced post-suppositions, which are constraints on output contexts, in contrast to presuppositions, which constrain input contexts.

Unlike presuppositions, post-suppositions are part of the proposal to update the Context Set / Common Ground (Stalnaker 1978) since they are part and parcel of regular truth conditions. Hence, both post-suppositional and at-issue meaning can be challenged, questioned etc. But post-suppositions are distinct from regular at-issue meaning with

[26]Contra van den Berg (1996) among others.

respect to their evaluation order: they constrain the final, global output context obtained after the regular at-issue update is interpreted.

But just as presuppositions (or implicatures in theories like Chierchia et al. 2009), post-suppositions can be satisfied / discharged non-globally, e.g., in the scope of distributivity operators.

Post-suppositions constrain quantificational alternatives (where a quantificational alternative is a context that is the result of interpreting a quantificational expression), not focus alternatives, as Krifka (1999) would have it. We therefore expect various *quantificational* operators (universals, modals, attitude verbs, negation etc.) to block the 'projection' of post-suppositions and discharge them locally, in their scope.

While similar to the actual (syntactic) wide scope that *bona fide* quantificational expressions can take, post-suppositional pseudo wide scope has slightly different semantic properties. These distinct properties enable us to account for the differences in semantic behavior between class A and class B modified numerals noticed in Nouwen (2010): class A modifiers contribute degree quantifiers that can take actual wide scope, while class B modifiers contribute post-suppositional cardinality requirements.

Class B modified numerals (e.g., *at most three books*) are maximal, just like cardinal definites (e.g., *the three books*). In Romanian, for example, both modified numerals and cardinal definites contain a definite article. But modified numerals and cardinal definites are different in two respects:[27] (*i*) numerals introduce a new maximal discourse referent / variable, while definites anaphorically retrieve an old discourse referent / variable that is presupposed to be maximal; (*ii*) the cardinality requirement contributed by numerals is a post-supposition, while the cardinality requirement contributed by definites is a presupposition.

The fact that modified numerals introduce a new variable storing a maximal set of entities and separately from this variable, a 'wide-scope' post-supposition makes them very similar to (discourse referents for) properties. This is very much in line with the property-based semantics for existential constructions in McNally (1998)[28] – and it might enable us to account for the fact that modified numerals, in contrast to cardinal definites, are felicitous in existential constructions, e.g., *There were* {*at most / *the*} *three books on the table* (from McNally 1998).

Ultimately, enriching contexts of evaluation with post-suppositions follows the same basic insight and strategy as the enrichment brought by classical dynamic semantics relative to static semantics. That is, enriching contexts of evaluation and (therefore) the inventory of opera-

[27] I am indebted to Sandy Chung, Donka Farkas, Jim McCloskey and Louise McNally for discussion of this point.

[28] I am indebted to Louise McNally for this observation (p.c.).

tors that can be defined over them enables us to keep our interpretation compositional and surface-based. The reason for this is that local operations over enriched contexts have global effects via the recursive definition of truth and satisfaction, which preserves these local contextual changes and non-locally passes them on.

We end the paper by mentioning three directions for future investigation. First, one should provide a compositional analysis for adverbial / 'floated' uses of *exactly*, *precisely*, *maximally*, *approximately* etc. in which the modifier is non-adjacent and can simultaneously target multiple numerals, as shown in the examples below.[29] Note, in particular, that the cumulative reading of (182) is one in which *at the most* simultaneously targets *four games* and *five days*.

(178) Three boys saw five movies, exactly / precisely / at (the) most.[30]

(179) It was a kind of pension where, at the most, there were four or five guests. (COCA)

(180) One person, at most, infects two others . . . (COCA)

(181) Everyone has to show me five dust balls, at least. (COCA)

(182) The league limits teams to playing two games in a row – or, at the most, four games in five days, NBA spokesman Tim Frank says. (COCA)

The analysis of modified numerals proposed in Brasoveanu (2010b) uses an unselective maximization operator for modified numerals, thereby analyzing them in parallel to adverbs of quantification like *usually*, *always* etc. (they both quantify over cases, in the sense of Lewis 1975). Maybe an analysis along these lines will be able to generalize to the phenomena exemplified above.

Second, it seems that the availability of cumulative vs distributive readings is sensitive to questions under discussion:[31] cumulative readings seem unavailable as answers to single *who*/*how many* questions like (183) below, but available as answers to multiple *wh*-questions like (184). Moreover, sentences with such cumulative readings seem to have a particular intonation pattern.

(183) How many boys saw exactly five movies?

(184) How many boys saw how many movies?

---

[29] I am indebted to Pranav Anand, Jim McCloskey and an anonymous Amsterdam Colloquium 2009 reviewer for discussion of this point.

[30] Interestingly, the adverbial modifier cannot be duplicated so that each occurrence can target a separate bare numeral: **Three boys exactly saw five movies exactly / precisely*.

[31] I am indebted to an anonymous Amsterdam Colloquium 2009 reviewer for bringing this point to my attention.

Cumulative readings are also unavailable in cases in which a parallelism discourse relation needs to be established, e.g.:

(185) Mary saw exactly five movies and exactly three boys did too / saw exactly five movies too.

This is not unexpected: quantifier scope, which also involves manipulating the evaluation order of certain expressions, is sensitive to questions under discussion and discourse relations. The observation that information structure and quantificational phenomena like cumulative readings are connected is already made in Zeevat (1994). The broader question is: how do focus alternatives and quantificational alternatives interact?

Finally, the relationship between a post-suppositional account of items like *exactly*, *precisely* etc. and their account as slack-regulating, halo-adjusting expressions along the lines of Lasersohn (1999) needs to be more closely investigated.[32]

Acknowledgments

I want to thank Lucas Champollion for providing the initial impulse to write this paper and Lucas Champollion, Jakub Dotlačil, three anonymous reviewers for *Journal of Semantics*, Rick Nouwen and Philippe Schlenker for many comments that greatly improved it. I am grateful to Pranav Anand, Donka Farkas, Jessica Rett, two anonymous reviewers for the Amsterdam Colloquium 2009 and the audiences of the UCSC S-Circle (Nov. 2009), Amsterdam Colloquium (Dec. 2009) and Stanford Scales Workshop (April 2010) for their detailed and helpful feedback on earlier versions of this paper and to Melissa Frazier and Grant McGuire for data discussion. The usual disclaimers apply.

References

van Benthem, J. (1986). *Essays in Logical Semantics*. Kluwer.

van den Berg, M. (1996). *Some Aspects of the Internal Structure of Discourse. The Dynamics of Nominal Anaphora*. PhD dissertation, University of Amsterdam.

Brasoveanu, A. (2007). *Structured Nominal and Modal Reference*. PhD dissertation, Rutgers University.

Brasoveanu, A. (2008). Donkey Pluralities: Plural Information States vs Non-Atomic Individuals. In *Linguistics and Philosophy* 31, 129-209.

Brasoveanu, A. (2010a). Decomposing Modal Quantification. In *Journal of Semantics* 27, 437-527.

[32] I am indebted to Pranav Anand for discussion of this point and to Benjamin Spector for outlining an analysis for *exactly* along these lines (p.c.).

Brasoveanu, A. (2010b). Modified Numerals as Post-suppositions. In *Logic, Language and Meaning: 17th Amsterdam Colloquium*, M. Aloni, H. Bastiaanse, T. de Jager & K. Schulz (eds.), Springer LNAI 6042, 203-212.

Brooks, P. & M. Braine (1996). What Do Children Know about the Universal Quantifiers *all* and *each*? In *Cognition* 60, 235-268.

Büring, D. (2008). The Least *at least* Can Do. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, C.B. Chang & h.J. Haynie (eds.), Cascadilla Proceedings Project, Somerville, MA, 114-120.

Champollion, L. (2010). Cumulative Readings of *Every* Do Not Provide Evidence for Events and Thematic roles. In *Logic, Language and Meaning: 17th Amsterdam Colloquium*, M. Aloni, H. Bastiaanse, T. de Jager & K. Schulz (eds.), Springer LNAI 6042, 213-222.

Chierchia, G., D. Fox & B. Spector (2009). The Grammatical View of Scalar Implicatures and the Relationship between Semantics and Pragmatics. To appear in *Handbook of Semantics*, P. Portner, C. Maienborn & K. von Heusinger (eds.), de Gruyter.

Constant, N. (2006). *English Rise-Fall-Rise: A study in the Semantics and Pragmatics of Intonation*. MA thesis, UC Santa Cruz.

Dekker, P. (2008). A Multi-dimensional Treatment of Quantification in Extraordinary English. In *Linguistics and Philosophy* 31, 101-127.

Dekker, P. (2010). *Dynamic Semantics*. University of Amsterdam ms. (to appear, Springer).

Dotlačil, J. (2010). *Anaphora and Distributivity. A study of* same, different, *reciprocals and* others. PhD dissertation, Utrecht University.

Evans, G. (1977). Pronouns, Quantifiers and Relative Clauses (I). In *The Journal of Canadian Philosophy* 7, 467-536.

Evans, G. (1980). Pronouns. In *Linguistic Inquiry* 11, 337-362.

Farkas, D.F. (2009). Varieties of Indefinites. In *Proceedings of SALT XII*, B. Jackson (ed.), CLC, Ithaca, NY, 59-83.

Ferreira, M. (2007). Scope Splitting and Cumulativity. In *Proceedings of the Workshop on Quantifier Modification* (ESSLLI 2007), R. Nouwen and J. Dotlačil (eds.).

Geurts, B. & R. Nouwen (2007). *At Least* et al: The Semantics of Scalar Modifiers. In *Language* 83, 533-559.

Gil, D. (1982). Quantifier Scope, Linguistic Variation and Natural Language Semantics. In *Linguistics and Philosophy* 5, 421-472.

Groenendijk, J. & M. Stokhof (1991). Dynamic Predicate Logic. In *Linguistics and Philosophy* 14, 39-100.

Hackl, M. (2000). *Comparative Quantifiers*. PhD dissertation, MIT.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*, PhD dissertation, University of Massachusetts, Amherst, published in 1988 by Garland, New York.

Heim, I. (2000). Degree Operators and Scope. In *Proceedings of SALT X*, Ithaca, NY, CLC Publications.

Karttunen, L. (1976). Discourse Referents. In *Syntax and Semantics Vol. 7*, J.D. McCawley (ed.), New York: Academic Press, 363-385.

Kratzer, A. (2000). *The Event Argument and the Semantics of Verbs*. UMass Amherst ms.

Krifka, M. (1986). *Nominalreferenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aspektklassen*. PhD dissertation, Universität München, published by Wilhelm Finck, Munich, 1989.

Krifka, M. (1999). At least Some Determiners Aren't Determiners. In *The Semantics/Pragmatics Interface from Different Points of View*, K. Turner (ed.), Elsevier, 257-291.

Krifka, M. (2007). More on the Difference between *more than two* and *at least three*. Paper presented at UC Santa Cruz, available at `http://amor.rz.hu-berlin.de/~h2816i3x/Talks/SantaCruz2007.pdf`.

Landman, F. (2000). *Events and Plurality*, Kluwer.

Lasersohn, P. (1999). Pragmatic Halos. In *Language* 75, 522-551.

Lauer, S. (2009). Free relatives with *-ever*: Meaning and Use. Stanford University ms.

Lewis, D. (1975). Adverbs of Quantification, in *Formal Semantics of Natural Language*, E. Keenan (ed.), Cambridge: Cambridge University Press, 3-15.

Link, G. (1983). The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach. In *Meaning, Use and Interpretation of Language*, R. Bauerle, C. Schwarze & A. von Stechow (eds.), Walter de Gruyter, Berlin, 302-323.

McNally, L. (1998). Existential Sentences without Existential Quantification. In *Linguistics and Philosophy* 21, 353-392.

Nouwen, R. (2003). *Plural pronominal anaphora in context*, PhD dissertation, UiL-OTS, Utrecht University, LOT Dissertation Series 84.

Nouwen, R. (2010). Two Kinds of Modified Numerals. In *Semantics and Pragmatics* 3: 1-41.

Rett, J. (2010). Equatives, Measure Phrases and NPIs. In *Logic, Language and Meaning: 17th Amsterdam Colloquium*, M. Aloni, H. Bastiaanse, T. de Jager & K. Schulz (eds.), Springer LNAI 6042, 364-373.

Robaldo, L. (2009). Distributivity, Collectivity and Cumulativity in terms of (In)dependence and Maximality. Univ. of Turin ms.

Roberts, C. (1987). *Modal Subordination, Anaphora and Distributivity*. PhD dissertation, UMass Amherst, published in 1990 by Garland, New York.

Roberts, C. (1989). Modal Subordination and Pronominal Anaphora in Discourse. In *Linguistics and Philosophy* 12, 683-721.

Schein, B. (1993). *Plurals and Events*. MIT Press.

Stalnaker, R. (1978). Assertion, in *Syntax and Semantics* 9, 315-332.

Sternefeld, W. (1998). Reciprocity and Cumulative Predication. *Natural Language Semantics* 6, 303-337.

Stone, M. (1999). Reference to Possible Worlds, RuCCS Report 49, Rutgers University.

de Swart, H. (1999). Indefinites between Predication and Reference. In *Proceedings of SALT IX*, CLC, T. Matthews & D.L. Strolovitch (eds.), 273-297.

Szabolcsi, A. (1997). Strategies for Scope Taking. In *Ways of Scope Taking*, A. Szabolcsi (ed.), Kluwer, 109-154.

Umbach, C. (2006). Why do modified numerals resist a referential interpretation? In *Proceedings of SALT XV*, CLC, 258-275.

Wang, L. (2005). *Dynamics of Plurality in Quantification and Anaphora*. PhD dissertation, UT Austin.

Zeevat, H. (1994). Applying an Exhaustivity Operator in Update Semantics. In *Ellipsis, Tense and Questions*, H. Kamp (ed.), DYANA-2 deliverable R 2.2.b, ILLC, Amsterdam.